Identification of novel functional domains of Rad52 in Saccharomyces cerevisiae

Plate, Iben

Publication date:
2006

Document Version
Publisher's PDF, also known as Version of record

Identification of novel functional domains of Rad52 in *Saccharomyces cerevisiae*

PhD thesis by Iben Plate

Center for Microbial Biotechnology
BioCentrum-DTU
The Technical University of Denmark
Spring 2006

Acknowledgements

This PhD work was conducted at the Center of Microbial Biotechnology (CMB) at the Technical University of Denmark from October 2002 – September 2005.

I would like to acknowledge my supervisor associate professor Uffe H. Mortensen for excellent supervision, support, discussions and inspiration. It has really been an educational journey with such an enthusiastic mentor! Members of the DNA repair group are also acknowledged for providing a positive working environment, both socially and scientifically.

I would like to thank Professor Patrick Sung at Yale Medical School for inviting me to work in his laboratory together with his talented research group. I will thank all members of his lab, in particular post doc Lumir Krejci for teaching me protein purification and how to survive working in "the cold room". I am very thankful to Wendy Bussen, Margaret Macris, Michael Sehorn, Hilarie Sehorn and Sander Granneman for making my stay at Yale so pleasant.

During my PhD I have had the pleasure of supervising some very gifted students; Swee Chuang Lim Jensen, Yao Shuo and Lifang Liu. The Chinese trio has successfully constructed *rad52* alleles deficient in repair center formation. Swee has continued on the repair center formation in her own PhD project and has contributed significantly with the fluorescent microscopy and *RAD51* overexpression experiments. Post doc Christian Müller is acknowledged for his work on analysing and quantifying the focus formation deficiency by fluorescent microscopy. Associate professor Michael Lisby at the University of Copenhagen is thanked for providing excellent pictures for the nuclear transport project. PhD student Line Albertsen has continued working on the nuclear transport of Rad52 in her own PhD project, and she is also thanked for the work she has done on the nuclear transport project by constructing mutations in NLS 1 and NLS 2. I would like to thank PhD student Gaëlle Lettier for always being helpful and for proofreading of the thesis. Likewise, Kiran Patil is thanked for helping me with setting up the thesis and reading and commenting the manuscript! I want to thank lab technicians Lene Christiansen and Mbarka Sersar for helping me in the laboratory. Also I want to thank PhD Tanja Thybo, and PhD Michael Lynge-Nielsen for practical advice and many helpful discussions.

I want to thank my office mates throughout my stay at CMB Christian, Steen, Margarida, Gaëlle, Jacob and Kiran for making every day at work enjoyable – also Mondays! And the beer club for making Fridays more fun! I want to thank everybody at CMB for providing such a nice atmosphere. Special thanks to the administrative staff in particular Birgitte Karsbøl for always being very helpful and the EDB departments Dan Aistrup for giving me trans-atlantic help and assistance when my computer crashed at Yale.

My PhD has been financed by Forskningsrådet and during the three years I have received several grants, which has made it possible to attend conferences and to fund my external stay in New Haven. I would therefore like to thank Familien Hede Nielsens Fond, Harboefonden, Civilingeniør Poul V. Andersens fond, Jørgen Esmers Mindelegat and Otto Mønsted-Fonden.

Preface

DNA double strand break repair (DSBR) is essential to maintain the integrity of the genome. Un-repaired or improperly repaired lesions can result in genomic instability, which in higher eukaryotes may lead to several diseases including cancer. Rad52 is an essential protein for homologous recombination and DNA DSBR in *Saccharomyces cerevisiae.* Among the genes involved in the homologous recombination and repair processes, *RAD52* deletion displays the most severe phenotypic defects in yeast. Moreover, the primary structure and several biochemical properties of Rad52 are evolutionary conserved from yeast to humans. Consequently, yeast is widely used as a model organism to study the function of Rad52. Not only the stable haploid state and availability of genetic tools make the yeast genome easy to manipulate, but also the sophisticated *in vivo* assays make it possible to monitor biological processes such as DNA repair and homologous recombination at the single cell level.

In vivo, DNA DSBs result in the accumulation of repair proteins in large repair complexes where multiple lesions are repaired simultaneously. Accordingly, along with other members of the *RAD52* epistasis group, Rad52 has been shown to relocalize from diffused nuclear distribution into concentrated repair centers when DNA DSBs are introduced. However, the molecular mechanisms underlying repair center formation are unknown. Thus, the primary aim of this study was to investigate how different sequence domains of Rad52 play a role in mobility-associated functionality.

First, it was investigated how ScRad52 is directed to the nucleus after translation, since no Nuclear Localization Signal (NLS) has previously been identified. For this purpose, a series of fluorescently tagged Rad52 truncation proteins were expressed and their cellular distribution monitored by fluorescent microscopy. The results from these experiments showed that the middle region of the Rad52 sequence is important for correct sorting since disruptions in this sequence resulted in cytosolic localization of the mutant protein. When this region was examined by sequence analysis a putative NLS was identified. Mutagenic disruption of this sequence motive resulted in cytosolic localization of the mutant protein strongly indicating that it is a functional NLS. Furthermore, when this mis-sorting mutant

protein was extended with the well characterized SV40 large T antigen NLS, it relocated to the nucleus and was biologically competent as cells expressing the NLS-tagged Rad52 protein survived DNA damage induced by an alkylating agent MMS. Together these results imply that a single NLS in Rad52 is required to ensure its nuclear transport and that transport is the only function of this sequence in Rad52.

Next, it was examined whether a specific region of Rad52 mediates the concentration of the protein into repair centers in response to DNA damage. To investigate this possibility, a large series of YFP-tagged *rad52* truncations and alanine substitution mutants were expressed and the formation of repair centers were monitored *in vivo* using fluorescence microscopy. In this study, several *rad52* alleles, which fail to concentrate into repair centers, were isolated. All mutations locate in the same region of Rad52 and one candidate mutation was selected for further characterization as strains expressing it showed the highest sensitivity to MMS. The mutant protein, which has a duster of four amino acid residues substituted with alanine residues, was unable to concentrate into repair centers spontaneously during S-phase as well as when DNA DSBs were introduced. The mutant displayed a severe phenotype including slow growth and MMS sensitivity, which indicate that the mutant strain lacks the ability of proper DNA DSB repair. The mutant protein was purified and characterized biochemically. None of the tested Rad52 functions such as DNA binding, ssDNA annealing and protein oligomerization were affected. Interestingly, the mutation is situated in a region of Rad52, which has previously investigated by using a mutant where this entire region of Rad52 has been eliminated in a truncation mutation. MMS sensitivity of strains expressing this truncation mutation can be fully alleviated if Rad51 is over-expressed. Interestingly, this is not the case for the alanine substitution mutant strain characterized in this thesis. In summary, a novel domain in Rad52 has been mapped during this study, which impacts the ability of Rad52 to participate in a repair center. Consequently, a model is proposed to explain how Rad52 and Rad51 collaborate to form a repair center.

Lastly, a biological characterization of a novel C-terminal DNA binding domain in Rad52 was performed. Recently, a mutant Rad52 protein composed only of the C-terminus of *RAD52* has been shown to be able to bind DNA and catalyze Rad51-mediated strand

exchange. In the study presented in this thesis, it was tested if this mutant protein also functions in DNA DSBR *in vivo*. The mutant protein was tagged with an NLS from SV40 large T antigen to ensure nuclear localization, and subsequently the NLS-tagged *rad52* mutant allele was investigated for its functionality in DNA DSBR by using MMS survival assays. Furthermore, its subcellular localization was monitored following DNA damage. A *rad52* strain expressing only the C-terminal DNA binding domain as well as one that expresses the C-terminal and the middle part of *rad52* showed a slight ability to suppress MMS sensitivity in a *rad52* null strain and to a larger degree in a truncated *rad52-D327* strain. However, when the mutant proteins were tagged with YFP and subjected to fluorescent microscopy they proved unable to form repair centers after DNA damage had been induced by MMS. Taken together, the results in this thesis imply the presence of an additional DNA binding domain in the C-terminus of Rad52, which plays a role in DNA DSBR not only *in vitro,* but also *in vivo*.

The work presented in this thesis shows, for the first time, the existence of three novel functional domains in ScRad52. The first domain contains an NLS that alone is responsible for the nuclear localization of ScRad52. This NLS is involved only in nuclear transport, and not in DNA repair. The second domain, which includes a stretch of four amino acid residues in the middle region of the Rad52 protein sequence, is responsible for the ability of Rad52 to concentrate into repair centers after introduction of DNA damage. The third domain characterized in this thesis is a C-terminal DNA binding region in Rad52, which is demonstrated to play a role in the repair of DNA DSBs. These three novel domains of Rad52 in *S. cerevisiae* provide further insight into the role of Rad52 and the molecular mechanisms underlying its role in DNA DSBR and repair center formation. Overall, the results not only highlight the importance of the multifunctional Rad52 protein, but also significantly contribute to the understanding of the DNA repair pathways that are highly important for maintaining genetic stability.

Extended Synopsis

Thesis motivation and aim

The living cell is constantly experiencing DNA lesions caused by exogenous as well as endogenous sources of DNA damage. The diverse lesions arise in the DNA from different causes; environmental agents such as the UV component of sunlight, ionizing radiation and genotoxic chemicals e.g. like those found in cigarette smoke. Products of normal cellular metabolism such as superoxide anions, hydroxyl radicals and hydrogen peroxide also pose a constant threat to the DNA integrity. The damage can result in a variety of lesions such as modified bases, mispaired bases, intrastrand crossbinding, interstrand crossbinding, pyrimidine dimers, single stranded breaks, and double stranded breaks. Consequently, evolution has molded a selection of sophisticated DNA repair systems that as a whole covers most of the insults inflicted on cells vital genetic information. Amongst these repair systems, the widely studied repair pathways include base excision repair, nucleotide excision repair, mismatch repair, non-homologous end-joining and homologous recombination.

Failure of these repair pathways, and consequent failure to repair DNA damage, may lead to a deleterious mutation rate, genetic instability, or even cell death. In higher eukaryotes, mutations in genes responsible for DNA repair and cell cycle regulation often lead to cancer development. Amongst DNA lesions, DNA double stranded breaks are very problematic as both strands of the DNA helix are affected. Recent discoveries that several human cancer-prone syndromes like Blooms syndrome, are caused by defects in the DSB repair have highlighted the importance of this repair pathway in maintaining genome integrity and avoiding cancer.

The growing number of genetic diseases related to defects in DNA repair mechanisms makes it essential to understand how these processes occur at the molecular level. Basic research in this field is much needed since there are still many black holes in the biology of DNA repair pathways and related processes, e.g. cancer development. Although, for years scientists all around the world have aimed at understanding the mechanisms underlying

cancer development, with approximately 200 cancer-related deaths per 100.000 persons in USA every year (NIH, 2004) the task is still highly relevant.

In yeast, DNA DSBs are mainly repaired by the homologous recombination pathway and to a lesser degree by the non-homologous end-joining pathway. Much of current understanding of the mechanism of DNA DSB repair is based on experiments performed in the yeast, *Saccharomyces cerevisiae*. *S. cerevisiae* is an ideal model organism to study DNA repair pathways. It is possible to knock out or modify relevant repair genes, study the mutant phenotype and monitor the cellular distribution of repair proteins *in vivo* with fluorescent probes fused to the genes of interest.

The *RAD52* epistasis group defines the main pathway for homologous recombination in eukaryotes. In yeast, Rad52 plays a fundamental role in the homologous recombination pathway and disruption of *RAD52* causes a severe recombination phenotype including extreme X-ray sensitivity, increased chromosome loss and failure to produce viable spores. Rad52 has been subjected to intensive studies over the last two decades, and knowledge on this key protein is still growing. Rad52 is a multi-domain protein of which several functional regions have been identified, including those that are involved in Rad52 self-association, in DNA binding and in permitting interactions with recombination proteins Rad51, RPA and Rad59.

The aim of this project was to map functional domains of Rad52 in *S. cerevisiae*, ScRad52. The focus was set on investigating three biological properties of ScRad52, namely, nuclear transport, repair center formation and DNA binding. The first aim was to identify a region of ScRad52 protein involved in nuclear sorting of the protein. Next aim was to explore how ScRad52 assembles in repair centers in response to DNA damage, and finally to assess the biological relevance of a recently identified DNA binding domain in ScRad52.

Mapping functional domains in ScRad52 is highly relevant for the understanding of the structure and function of human Rad52. Several biochemical properties of the Rad52 protein as well as the homologous recombination pathway are evolutionary conserved.

Thus, the growing knowledge of ScRad52 function can play an important role in elucidating the repair mechanisms that are often involved in cancer development in humans.

Summary of results

The function of the Rad52 protein has been studied intensively and domains for specific functions have been assigned. Accordingly, the N-terminal contains DNA binding function and Rad52- and Rad59 association function whereas the C-terminal mediates Rad51 association as well as Rad52 association. However, several questions regarding the molecular functionality of Rad52 still remain open. The first question is how the Rad52 protein is transported from the cytosol to the nucleus following translation? Next question is how Rad52 is concentrating into large repair centers when the cell experiences DNA DSBs? And last question to be asked in this study is if there is a biological function of the DNA binding domain recently identified in the C-terminal of Rad52?

Before the beginning of this PhD project in the fall of 2002, no putative NLS had been identified in Rad52. Several research groups had assigned regions of the protein to transport, but no experimental data was available to confirm these assumptions. In this study, a large series of Rad52 mutant proteins tagged with fluorescent probes were constructed to identify a region in the protein involved in the nuclear transport. A cluster of positively charged amino acid residues located in the middle of the protein sequence was found to be indispensable for nuclear localization of Rad52. It was proposed that this stretch of amino acid residues makes up an NLS in Rad52 that is responsible for the nuclear localization of the Rad52 protein. This proposed NLS is positioned in a part of Rad52, which is not widely evolutionary conserved, which is in good agreement with other observations made in this study. Fluorescent microscopy showed that Rad52 from the yeast species *K. lactis* tagged with YFP can sort to the nucleus when expressed in a *S. cerevisiae rad52D* strain whereas Rad52 from *M. musculus* can not. This suggests that the nuclear transport is not conserved throughout evolution, but is (at least) partly species specific.

One of the evolutionary conserved properties of Rad52 is that the protein relocates in response to DNA damage. Following DNA damage the Rad52 protein concentrates into repair centers consisting of hundreds to thousands of Rad52 molecules, which is otherwise evenly distributed in the nucleus. These repair centers are believed to be the

site of ongoing repair of DNA DSBs, and the repair proteins organize in multi-protein complexes in response to DNA damage by sequential recruitment. Still, it is unclear what mechanisms drive Rad52 into these large repair complexes in response to DNA damage. In this research project, a systematic mutational study of Rad52 was made to learn more about what makes Rad52 concentrates in response to DNA damage. The aim was to identify a region in Rad52 responsible for forming repair foci, and a region in the middle of the Rad52 sequence was successfully identified. When four amino acid residues were substituted with alanine in this region, the resulting protein was unable to be incorporated into a repair center. The mutation resulted in a severe phenotype and the strain was highly sensitive to MMS-induced damage and grew slowly. These results show that the repair centers are essential for efficient repair of DNA DSBs. Although the mutant protein was defective in focus formation, none of the other tested Rad52 activities, namely DNA binding, ssDNA annealing, and oligomerization, were affected. The mutant Rad52 protein also associated physically with Rad51, but it was not possible to establish an assay to successfully detect interaction to RPA.

The ss- and dsDNA binding property of Rad52 is believed to be central for correct DNA DSBR. Recently, members of Professor Patrick Sung's laboratory at Yale University have identified a novel DNA binding domain in the C-terminal of Rad52. However, the function of this additional DNA binding domain is unknown and it is also not clear why the Rad52 protein has two distinct DNA binding domains. Do these two domains have different functions in the DNA repair process? *In vitro*, Rad52 species expressing the C-terminal domain was capable of annealing complementary ssDNA and mediate Rad51-catalyzed strand exchange. In this study, the biological relevance of this region was characterized *in vivo* using survival assays together with fluorescent microscopy. A *rad52* deficient and a *rad52* truncated strain were partly suppressed in MMS-sensitivity assays when co-expressed with the C-terminal of Rad52 and with the C- as well as the middle part of Rad52. This indicates that the C-terminal DNA binding domain plays a role in the repair of DNA DSBs. The mutant proteins were tagged with fluorescent labels and their cellular localization were monitored as well as their ability to form repair centers in response to DNA damage. The results showed however, that the mutant proteins were unable to concentrate into repair centers visible with fluorescent microscopy.

In summary, this PhD work contributes to the field of DNA repair in identifying and characterizing novel functional domains to Rad52; and also shows how fluorescence microscopy can be used to visualize biological functions *in vivo* when monitoring fluorescently tagged repair proteins.

In the next section the players in DNA DSBR, in particular homologous recombination, will be introduced. After that, the three studies conducted in this PhD work namely nuclear transport of Rad52, repair center formation of Rad52 and biological characterization of a C-terminal DNA binding domain, will be described. The three studies are to be published soon and therefore it was decided to write the three chapters in a manuscript-like format.

Contents

1. Introduction

In this section, the repair pathways homologous recombination (HR) and non-homologous end-joining (NHEJ) will be introduced briefly. After this short discussion, two models for DNA DSB repair by HR will be described. Next, some of the most important proteins involved in the HR repair pathway and their interconnections will briefly be presented followed by an individual description of the proteins and their role in DNA DSBR. The description of the HR pathway is followed by a brief introduction to the assays used to study DNA DSBR in yeast. Lastly, background for the three projects of this PhD study involving nuclear transport of Rad52, formation of Rad52 repair centers as well as DNA binding of Rad52 will be introduced.

DNA DSBs occur spontaneously or as a consequence of exposure to DNA damaging agents. When a cell incurs a DNA DSB it decides if the break should be repaired from a homologous sequence by HR or by the NHEJ pathway. It is still uncertain how the cell decides to use one repair pathway over the other and channels the DSB into a certain pathway. In NHEJ, the DNA ends are aligned and rejoined using little or no sequence homology and in HR the two broken DNA ends are repaired using information from a homologous sequence such as the sister chromatid or the homologous chromosome. The two pathways NHEJ and HR have been reviewed in recent years and will be introduced below (Paques and Haber, 1999), (Symington, 2002), (Hefferin and Tomkinson, 2005).

1.1 NHEJ

The NHEJ pathway repairs a DNA DSB by ligating the two broken ends together and does not require homology between the two molecules. End-joining events are mutagenic since they can lead to small DNA deletions as the broken ends are processed prior to the religation.

Figure 1 shows a conceptual model of NHEJ. NHEJ has been conserved during evolution, and many proteins involved in the pathway are the same in mammals and in yeast.

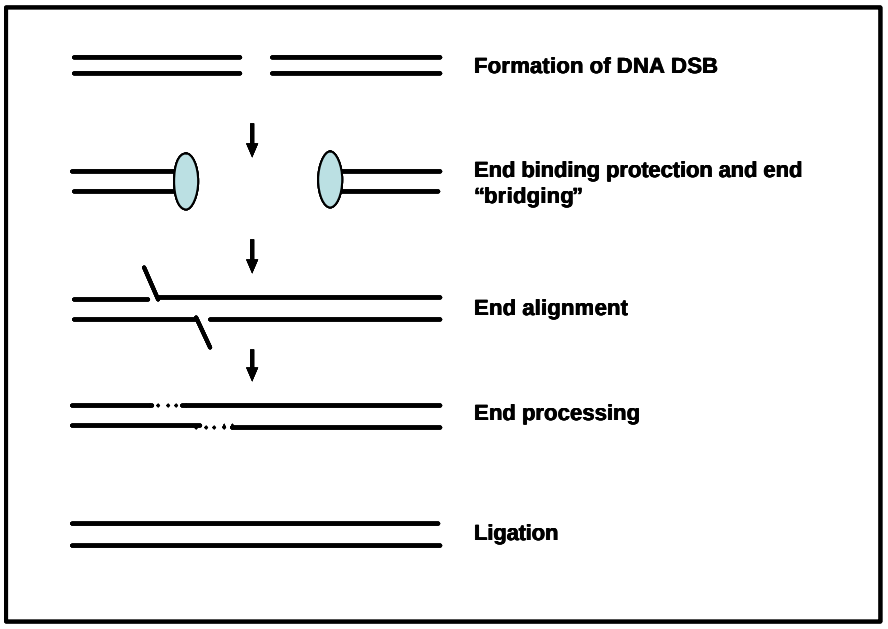


Figure 1. Schematic representation of the non-homologous end-joining pathway.

Following the break, specific proteins bind the two DNA ends to prevent nucleolytic degradation. Next, the breaks are held together in close proximity by a "bridging" mechanism (Hefferin and Tomkinson, 2005). Chen *et al*. suggest that the bridging in *S. cerevisiae* serves as a scaffold on which the NHEJ machinery is assembled (Chen *et al.*, 2001). The end-bridging event leads to cell cycle arrest until the lesion has been repaired. Furthermore, at this point NHEJ repair proteins necessary for processing and ligating the DNA ends are recruited. The ends are held together so they can be processed before religation. A DNA DSB usually results in DNA ends that are not directly suitable for ligation, and consequently they need processing prior to ligation. The next step in the process is to get the two DNA ends aligned. This can be facilitated by short complementary sequences of 1 - 4 nucleotides. This micro-homology depending alignment can result in mismatched flaps or gaps that are processed by nucleases and polymerases. When the DNA ends have been properly processed and aligned, DNA ligases are recruited to catalyze the ligation of the two strands. Due to possible processing of the DNA ends the resulting repaired DNA molecule can have DNA alterations (Paques and Haber, 1999), (Symington, 2002), (Hefferin and Tomkinson, 2005).

1.2 HR

Several models for DNA DSB repair by homologous recombination have been proposed. These models include the double strand break repair (DSBR) model, the synthesis

dependent strand annealing (SDSA) model, the break induced replication (BIR) model and the single strand annealing (SSA) model. Two of these models, DSBR and SDSA, will be described in the following.

1.2.1 A DSBR model

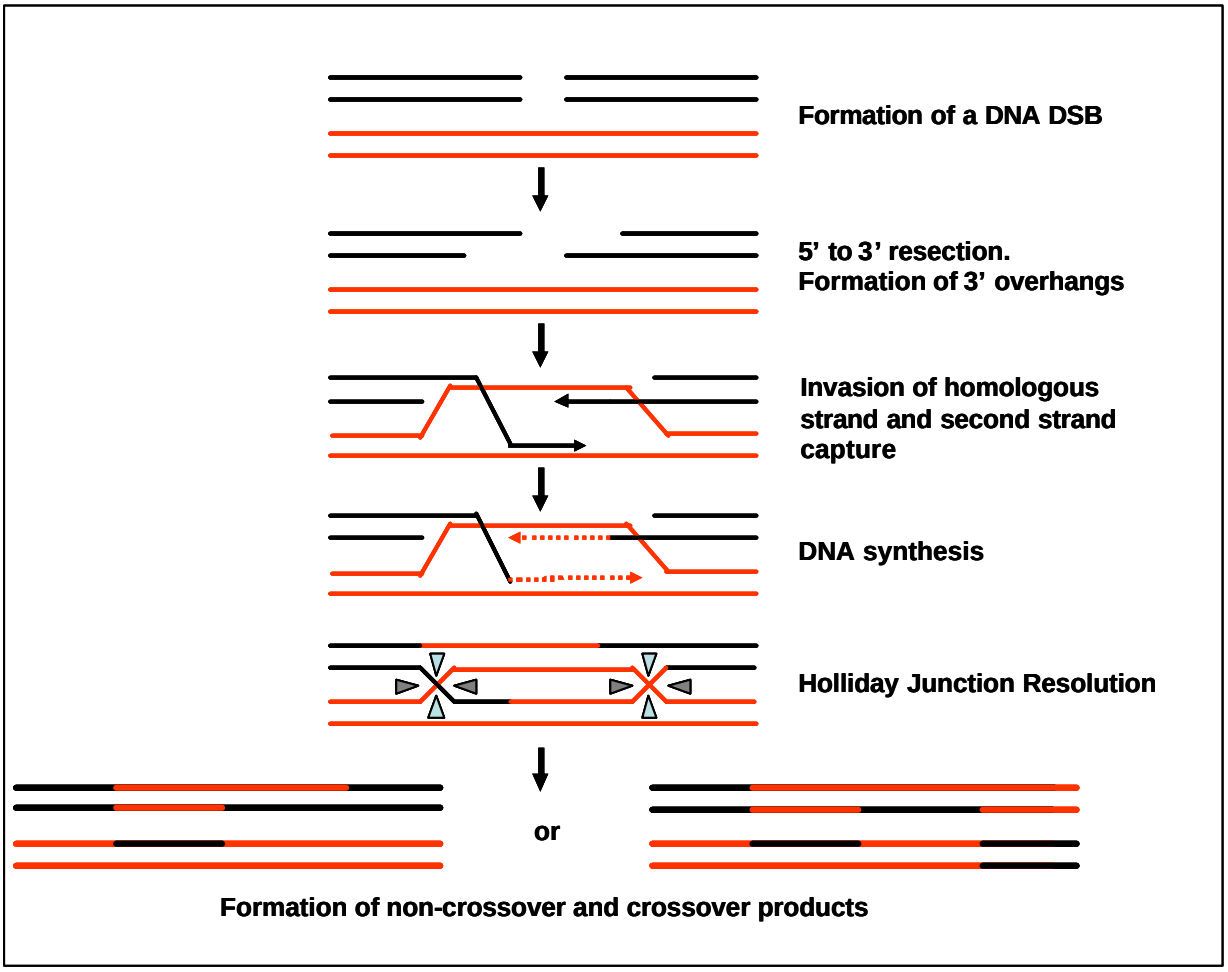


Figure 2. Schematic representation of a DSBR model.

DNA DSBs can be repaired by homologous recombination. The DSB lesions either occur spontaneously or are induced by DNA damaging agents. Following the break, the ends are resected in 5' – 3' polarity creating 3' single stranded tails. The 3' tail is recombinogenic and invades the homologous duplex, promotes DNA synthesis and forms a D-loop. The D-loop will pair with the non-invading strand (second end capture) that is also extended by DNA synthesis. As a result, two four-stranded intermediates, double Holliday junctions, are formed (figure 2). The structures can be resolved in two different orientations resulting in

either crossover or non-crossover products. If the structures were resolved randomly, it would result in equal numbers of crossover and non-crossover products. However, the formation of non-crossover products in gene conversion is low, and to account for this the SDSA recombination model was proposed.

1.2.2 The SDSA model

In SDSA, the initial steps are the same as in the DSBR model. One of the overhanging 3' tails invades a homologous sequence, such as the sister chromatid or homologous chromosome, and primes new DNA synthesis. The invading DNA strand is released, either as it is synthesized, or following a short round of DNA synthesis, which leaves a long 3' single strand. This single strand can anneal with the single strand at the other end of the break. The result is gene conversion that is not associated with a crossover event (figure 3).

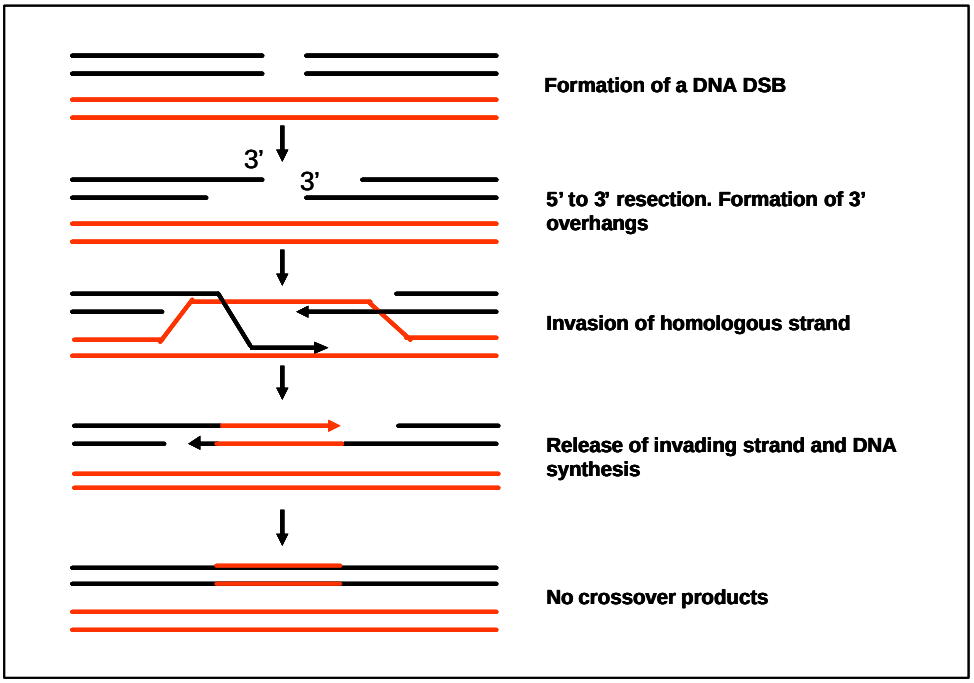


Figure 3. Schematic representation of a SDSA repair model.

1.3 Proteins involved in HR repair

The pathway for the repair by HR involves several proteins. Since protein-protein interactions, DNA binding properties and structural characteristics of proteins are all believed to play a role in DNA DSB repair it is important to define the relationships among

them. In the following, the known key proteins participating in repair of DNA DSB by HR will be introduced. The schematic drawing of the known physical interactions between some of the DNA repair proteins in yeast cells is depicted below (figure 4). The drawing shows the complexity of interaction among the proteins involved and is generated using the MINT database and illustrates physical protein-protein interactions based on experimental yeast two-hybrid and immunoprecipitation experiments. In the following, the relevant proteins in DNA DSBR will be described individually followed by a description of the DNA DSBR model including the proteins involved.

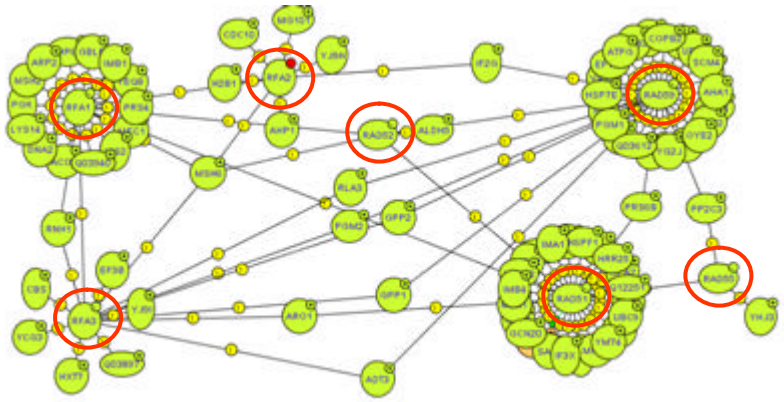


Figure 4. Protein-protein interaction among some of the proteins involved in DNA DSBR using the MINT database. The plus signs indicate that the particular protein is involved in additional interactions. The numers 1, 2, 3 etc. refers to the number of references the interaction is based on.

When the DNA strand encounters a lesion, a sophisticated system is set into play. The damage is sensed, signals send to the repair proteins and the lesion is repaired. The cellular response to DNA damage can be seen as a classical signaling-transduction cascade with DSB sensors, transducers and effectors (Jackson, 2002). The “signal” is the DNA damage, DNA binding proteins act as “sensors” of the break and activate the “tranducers” that again activate the “effectors”. The focus of this thesis is on one of the “effectors”; the DNA repair process. Thus, the attention is not put on the sensors of the DSBs or the transducers following a lesion, but is mainly put on the actual players in the DSB repair process. In the following, some of the most important proteins involved in the DNA DSBR pathway will be introduced and their role in the pathway discussed.

1.4.1 Mre11/Rad50/Xrs2 complex

The 83 kDa Mre11 protein is involved in genome surveillance, and is the first protein detected at the break site. Mre11 and the ATM-related Tel1 kinase appear before the DNA strand is resected (Lisby *et al.*, 2004). Mre11 is highly conserved among eukaryotes and is the homologue of human ATR. Mre11 forms a complex with Rad50 and *RAD50* and *MRE11* are required for the nucleolytic processing of DNA DSB in yeast (Trujillo and Sung, 2001). In yeast, Rad50 is 153 kDa and contains a nucleotide binding domain, and two coiled-coil domains (Alani *et al.*, 1989). The Rad50 protein is evolutionary conserved and the two coiled-coil domains are flanked by typical Walker A and Walker B ATPase motifs (Alani *et al.*, 1989). Trujillo and Sung purified the *S. cerevisiae* Rad50/Mre11 complex. Like Mre11, the Rad50/Mre11 complex exhibit a 3' to 5' exonuclease activity with preference for duplex DNA ends (Trujillo and Sung, 2001). Unlike human Rad50, the yeast homolog does not stimulate the exonuclease activity of Mre11 (Trujillo and Sung, 2001), (Paull and Gellert, 1999) and Rad50 alone does not have any exonuclease activity (Trujillo and Sung, 2001). Mre11 and the Rad50/Mre11 complex comprise not only exonuclease, but also have endonuclease activity for specific DNA structures (Lewis *et al.*, 2004), (Krogh *et al.*, 2005). Using scanning force microscopy de Jager *et al*. show that human Rad50 and Mre11 form a flexible complex that tethers DNA ends (de *et al.*, 2001). The complex consists of two Rad50 and two Mre11 molecules and the architecture is a globular domain from which two rod-like coiled-coil structures protrude (de *et al.*, 2001), (Anderson *et al.*, 2001). Mre11 binds as a dimer between the catalytic domains of Rad50 and brings together the catalytic domain of Mre11 and ATPase and DNA binding domains of Rad50 (Anderson *et al.*, 2001) (Figure 5).

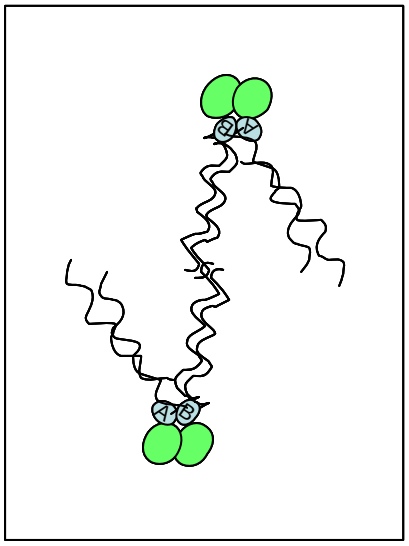


Figure 5. Mre11/Rad50 complexes connected by a "hook" domain. The Rad50 N-terminal Walker A (A) and Walker B (B) motifs are in close proximity with two Mre11 monomers (in green). The flexible Rad50 arms are protruding from the central core and proposed to interact through a "hook" domain (Lichten, 2005).

Although the “arms” are very flexible, in the presence of DNA they undergo a conformational change and become parallel (Moreno-Herrero *et al.*, 2005) and capable of tethering DNA. Different DNA molecules are held together by the “arms” of the complex, with preferential binding to linear DNA (de *et al.*, 2001). The biological relevance of this structure could be to keep the DNA ends in close proximity for further processing (de *et al.*, 2001) or to hold the DNA ends in close proximity to facilitate the search for homology (Anderson *et al.*, 2001).

Rad50 and Mre11 are associated with Xrs2 to form a complex, the MRX complex (Usui *et al.*, 1998), (Johzuka and Ogawa, 1995). Xrs2 is a functional homologue of mammalian Nbs1. Though the two proteins are thought to be functional homologues, they only share limited sequence similarity in the N-terminal. The function of Xrs2 in yeast and Nbs1 in mammals is not fully elucidated, but Nbs1 stimulates the unwinding a DNA duplex and cleavage of fully paired hairpin by human Mre11 and Rad50 (Paull and Gellert, 1999). A unique role of Xrs2 has recently been suggested by Trujillo *et al*. that found Xrs2 to be involved in the interaction of the MRX complex with DNA ends by targeting Rad50 and Mre11 to DNA ends (Trujillo *et al.*, 2003). This targeting to DNA might be explained by the observation that Xrs2 binds preferentially to ssDNA secondary structures.

The MRX complex is recruited to DSBs before 5’ resection exposes ssDNA, and the complex has been suggested to be involved in recognizing DNA double strand breaks (Lisby *et al.*, 2004). Yeast *mre11D*, *rad50D* and *xrs2D* strains have slow resection of HO-induced DNA DSBs at the MAT locus, indicating a role for the MRX complex in processing the ssDNA ends (Ivanov *et al.*, 1994) (Tsubouchi and Ogawa, 1998). Following a DNA DSB, the DNA ends are processed in 5’ to 3’ polarity, which obviously excludes the 3’ to 5’ exonuclease Mre11 as responsible for the resection alone. Trujillo and Sung proposed that the structure-specific activities of Rad50/Mre11 together with a yet unidentified DNA helicase may play an important role in processing the DNA DSB (Trujillo and Sung, 2001).

1.4.2 Sae2

Sae2 is associated with the MRX complex and has been proposed to play a role in the resection process (McKee and Kleckner, 1997). Sae2 has no known homologues in other

species and the function of the protein in recombination is still not fully understood. However, an intermediary role of Sae2 has been suggested. Mre11-GFP foci disassemble before Rad52 foci form at the site of the break, and are turned over slower in *sae2D* cells than in wild type Sae2 cells. Furthermore, the appearance of Rad52 foci is delayed in *sae2D* cells. As a support to the hypothesis that Sae2 is acting as a co-factor, is another observation that Sae2 foci appear at the same time as Mre11 foci disappears and Rad52 foci appears (Lisby *et al.*, 2004). Work by Clerici and co-workers suggest a role for Sae2 in both resection and bridging of the DNA DBSs. They observe that *sae2D* deficient cells have a slow resection of HO-induced DNA DSBs and are severely reduced in repair of HO-induced DSB by the single strand annealing pathway (Clerici *et al.*, 2005).

1.4.3 RPA

An important protein comes into play when the DNA ends have been resected. Replication protein A, RPA, was first identified as an ssDNA binding protein important for *in vitro* replication of Simian virus 40 (Fairman and Stillman, 1988). In yeast, RPA forms a heterotrimer consisting of three subunits of 70, 34 and 14 kDa, encoded by the *RFA1*, *RFA2* and *RFA3* genes (Brill and Stillman, 1989), (Alani *et al.*, 1992), (Wold, 1997). The homology between the two largest subunits of yeast and human RPA is high with more than 40% identity in amino acid sequence (Heyer *et al.*, 1990). RPA is an evolutionary conserved protein involved in several functions such as DNA replication and homologous recombination (Erdile *et al.*, 1991), (Alani *et al.*, 1992), (Firmenich *et al.*, 1995). Deletion of any of the three subunits of RPA in yeast leads to lethality (Heyer *et al.*, 1990).

Brill and Stillman and later Alani and co-workers performed a characterization of the DNA binding properties of yeast RPA and found RPA to bind ssDNA with a high affinity and co-operatively. The protein bound as a 1:1:1 complex of the three subunits of RPA (Brill and Stillman, 1989), (Alani *et al.*, 1992). The human RPA bind ssDNA via the large 70 kDa subunit, and work by Pfuetzner *et al*. propose that the smallest fragment of RPA1 harboring ssDNA binding activity comprises 181-422 amino acid residues of the total 441 (Pfuetzner *et al.*, 1997).

Additionally, RPA removes secondary structures from ssDNA and stimulate the RecA strand exchange reaction (Brill and Stillman, 1989), (Alani *et al.*, 1992). Later, it was demonstrated how RPA stimulated the Rad51-mediated strand exchange reaction provided that RPA is added to a preexisting complex of Rad51 and ssDNA. If it is added to the ssDNA before Rad51 it will inhibit the reaction. The result suggests that the two proteins compete for the same binding site on the ssDNA (Sung, 1994), (Baumann *et al.*, 1996), (Sugiyama *et al.*, 1997), (Shinohara and Ogawa, 1998), (New *et al.*, 1998). Additionally, it has been proposed that RPA exerts its effect on the presynapic complex by removing secondary structures on the ssDNA that Rad51 binds (Sugiyama and Kowalczykowski, 2002).

RPA interacts specifically with Rad52, and this interaction is crucial in the DNA repair process. This association will be described in more detail in the Rad52 section.

1.4.4 Rad51

Yeast *RAD51* encodes a 43 kDa protein with a Walker A and B motif for ATP binding and hydrolysis (Aboussekhra *et al.*, 1992), (Basile *et al.*, 1992), (Shinohara *et al.*, 1992). *RAD51* is a *recA* homolog and the sequence is conserved in organisms like the alga *Chlamydomonas reinhardtii*, yeast, mouse and human (Shalguev *et al.*, 2005), (Shinohara *et al.*, 1992), (Morita *et al.*, 1993). In yeast, *rad51*? mutants display only a low decrease in mitotic recombination between inverted repeats, but high sensitivity to ionizing radiation and meiotic inviability (Bai and Symington, 1996). The phenotype in vertebrates is more severe as deletion of *RAD51* results in early embryonic death in mice (Lim and Hasty, 1996).

Rad51 associates with itself, Rad52 (Milne and Weaver, 1993), (Shen *et al.*, 1996a), (Shinohara *et al.*, 1998), Rad54, (Johnson and Symington, 1995), (Jiang *et al.*, 1996), (Clever *et al.*, 1999) and Rad55 (Hays *et al.*, 1995). Rad51 catalyzes homologous DNA pairing between homologous single- and double-stranded DNA substrates and catalyzes strand exchange in biochemical assays (Sung, 1994). Yeast Rad51 forms right-handed filaments on ssDNA (Sung and Robberson, 1995) and dsDNA (Ogawa *et al.*, 1993). The Rad51 filament on dsDNA is biologically inactive in mediating the homologous DNA pairing

and strand exchange, whereas the filament on ssDNA mediates pairing and strand exchange (Sung and Robberson, 1995). The filament assembly is facilitated by RPA that removes secondary structures in the DNA (Sung, 1994), (Sugiyama *et al.*, 1998). This occurs when RPA is added after Rad51 has been given an opportunity to nucleate onto the ssDNA template. If RPA instead is added at the same time as Rad51 and the two proteins compete for binding to the ssDNA, the filament assembly is decreased and the efficiency of strand exchange is reduced (Sung, 1997b). If RPA is allowed to bind ssDNA before Rad51, it inhibits the strand exchange reaction (Sugiyama *et al.*, 1997). Rad52 can alleviate the inhibitory effect of RPA and is proposed as a co-factor in the Rad51-catalyzed strand exchange reaction (Sung, 1997a). The order of which Rad51 and Rad52 arrive at the break site is a subject of controversy. More details on this and interaction between Rad51 and Rad52 will follow in the section on Rad52 protein.

1.4.5 Rad59

Yeast *RAD59* was discovered through its requirement in *RAD51*-independent spontaneous mitotic recombination between inverted repeats (Bai and Symington, 1996). Rad59 is required for single strand annealing and for efficient DSB induced gene conversion (Sugawara *et al.*, 2000). *RAD59* encodes a 238 amino acid protein and is homologous to the N-terminal region of Rad52 (Bai and Symington, 1996). Rad52 and Rad59 have some overlapping functions. Like Rad52, Rad59 binds to ss- and dsDNA and promotes annealing of complementary single-stranded DNA. However, this annealing is not stimulated by RPA like it is for Rad52 (Petukhova *et al.*, 1999). In addition, work by Davis and Symington suggested that Rad59 is capable of self-association (Davis and Symington, 2003). Although Rad59 shares some properties with Rad52, it is unable to displace RPA to facilitate the loading of Rad51 (Bai *et al.*, 1999), (Davis and Symington, 2001). Bai and Symington propose that the function of Rad59 is to enhance the activity of Rad52 in strand annealing in repair pathways like gene conversion, break induced replication and single strand annealing (Bai and Symington, 1996). It has also been suggested that the role of Rad59 in single strand annealing is to either anneal strands or stabilize the annealed strands. In the gene conversion reaction, Sugawara and co-workers propose that Rad59 acts by catalyzing the formation of a paranemic joint structure, which can then be converted into a plectonemic joint structure (Sugawara *et al.*, 2000). A more

direct role of Rad59 has also been introduced by Jablonovich *et al*. that find *RAD59* to be responsible for a *RAD52*-dependent, *RAD51*-independent type of ectopic gene conversion. They propose that in the direct repeat recombination, Rad59 participates in a mechanism that can promote recombination in the absence of Rad51 filaments. However, it was unable to replace the Rad51 filaments in UV-induced gene conversion (Jablonovich *et al.*, 1999).

1.4.6 Rad54

Rad54 and the Rad54 homologue Rdh54 show properties that are characteristic of DNA helicases. Rad54 has dsDNA-dependent ATPase activity and promotes a conformational change of closed-circular duplex due to the creation of positive and negative writhe. Rad54 is believed to stimulate the Rad51 filament in the strand exchange assay (Mazin *et al.*, 2003), (Solinger *et al.*, 2001), (Solinger and Heyer, 2001). Furhtermore, Rdh54 promotes D-loop formation with Rad51 (Petukhova *et al.*, 2000).

1.4.7 Rad55/Rad57

The *RAD55* and *RAD57* genes of *S. cerevisiae* encode proteins with sequence similarity to RecA and Rad51 and are considered to be Rad51 paralogs. The Rad55 and the Rad57 proteins form a stable heterodimer. Both Rad55 and Rad57 proteins enhance *in vitro* strand exchange. Genetic data indicates that they participate directly in *in vivo* strand exchange, playing a supporting role that is sometimes dispensable (Fortin and Symington, 2002).

1.5 Rad52

Since Rad52 is the main focus of this thesis, the following description of Rad52 is more thorough than for the proteins discussed above.

Screens for radiation-sensitive mutants in budding yeast led to the discovery of the *RAD52* epistasis group genes that define the main pathway for homologous recombination in eukaryotes. Accordingly, the Rad52 protein was first isolated in a genetic screen of radiation sensitive mutants induced by nitrous acid in yeast *S. cerevisiae* (Resnick, 1969).

The *rad52-1* (A90V) mutant was isolated as a strain highly sensitive to X-ray. The mutation conferred low sporulation efficiency and spore viability (Resnick, 1969). A few years later another *rad52* mutation *rad52-2* (P64L) was isolated in a similar screen (Snow, 1967). The mutant strain was also very sensitive to ionizing radiation (Game and Mortimer, 1974). Since then *rad52-1* and *rad52-2* have been subject of intensive studies for understanding Rad52 protein function.

1.5.1 *RAD52* is evolutionarily conserved

RAD52 was first cloned in *S. cerevisiae* in 1984 (Adzuma *et al.*, 1984) and has since been isolated from various organisms like fungi, fish, chicken, mouse and human (Sakuraba *et al.*, 2000), (van den *et al.*, 2001), (Takahashi and Dawid, 2005), (Bezzubova *et al.*, 1993), (Bendixen *et al.*, 1994), (Muris *et al.*, 1994), (Shen *et al.*, 1995), (Park, 1995). Although no homologues of Rad52 have been identified in prokaryotes, the Rad52 protein shares structural and functional similarities with a class of recombination proteins found in bacteria and phage.

The ORF of *S. cerevisiae RAD52* encodes a 505 amino acid residue protein (Adzuma *et al.*, 1984). However, it has been shown that transcription of *RAD52* produces multiple protein species because *RAD52* has five putative start sites. *RAD52* is expressed from the third, fourth and fifth start site and each of the resulting proteins are competent in DNA repair (de Mayolo – submitted). With no transcription from the first or the second start codon, the third start codon is the first potential start codon and expression from this start codon of *RAD52* results in a 471 amino acid protein with a predicted molecular weight of 56 kDa (Shinohara *et al.*, 1998). In this thesis, the old numbering is used when referring to Rad52 amino acid residues to avoid confusion.

The N-terminal of Rad52 is evolutionary conserved, in particular the region spanning amino acid residues 34 -198 (Mortensen *et al.*, 2002) (figure 6). The figure shows a functional domain map of Rad52 protein, which will be explained later. On the figure, the percent identity with mouse Rad52 is also shown, with the dark color indicating the most conserved regions.

.

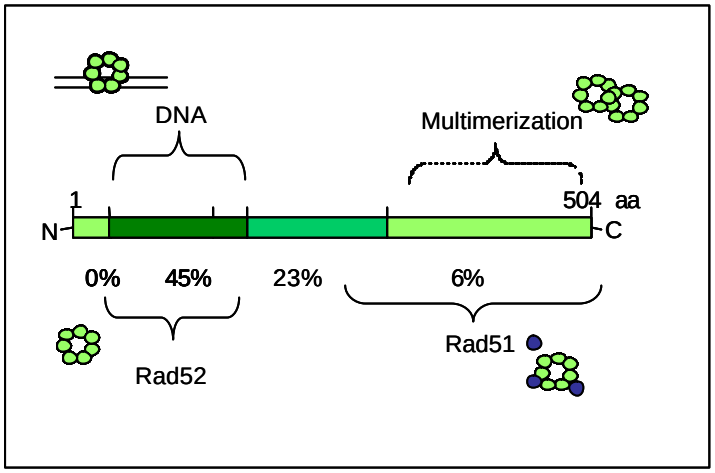


Figure 6. Functional domain map of Rad52. The N-terminal is conserved, indicated by the dark color. The numbers below the figure shows percentage identity with mouse Rad52. Domains involved in DNA binding, Rad51 binding, self-association, and multimer formation are also placed on the model.

1.5.2 Rad52 crystal structure

Two research groups have crystallized the N-terminal of human Rad52 simultaneously. Kagawa and co-workers have determined the crystal structure of the N-terminal 212 amino acid residues, Rad52-1-212Δ (Kagawa *et al.*, 2002) and Singleton *et al*. crystallized the first 209 amino acid residues, Rad52-1-209Δ (Singleton *et al.*, 2002). Both crystal structures of Rad52 show the protein in a closed ring consisting of eleven monomers.

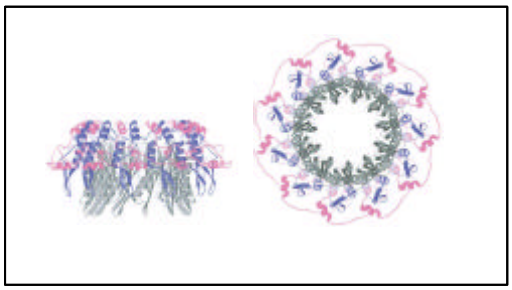


Figure 7. Crystal structure of N-terminal human Rad52. Picture from Kagawa *et al.* shows the side view and the bottom view of the predicted structure of the undecameric Rad52-1-212Δ ring. Taken from (Kagawa *et al.*, 2002).

1.5.3 Rad52 binds to DNA

- ssDNA and dsDNA

Several biochemical properties of Rad52 are evolutionarily conserved, hence both yeast and human Rad52 bind ss- and dsDNA (Mortensen *et al.*, 1996), (Shinohara *et al.*, 1998)

(Van *et al.*, 1998). Ranatunga and co-workers have shown that the Rad52 mutant protein expressing amino acid residues 1 – 192 binds ssDNA to the same extend as wild type Rad52, suggesting that the DNA binding domain of Rad52 is contained within the N-terminal part of the protein (Mortensen *et al.*, 1996), (Ranatunga *et al.*, 2001). Shinohara *et al.* used labelled FX174 ss- and dsDNA in binding assays with yeast Rad52 and found Rad52 to bind both substrates, with a higher affinity for ssDNA than for dsDNA (Shinohara *et al.*, 1998).

- DNA annealing

Several laboratories have shown that Rad52 promotes the annealing of complementary DNA strands *in vitro* (Mortensen *et al.*, 1996), (Kagawa *et al.*, 2001). Van Dyck *et al*. have visualized the annealing reaction mediated by human Rad52 by using electron microscopy and found Rad52 to anneal complementary resected single stranded tails and that these annealing products remained bound by the Rad52 protein (Van Dyck *et al.*, 2001). In addition, Shinohara and co-workers showed that RPA stimulates the annealing by Rad52 in a concentration-dependent manner. They proposed that RPA recruits Rad52 onto the ssDNA and stimulates the binding of Rad52 to ssDNA. The annealing activity of yeast Rad52 is specifically stimulated by RPA and not by *E. coli* SSB or T4 phage gp32 (Sugiyama *et al.*, 1997), (Shinohara *et al.*, 1998).

1.5.4 Rad52 binds to itself

Human Rad52 has been shown to self-associate both *in vivo* and *in vitro* (Shen *et al.*, 1996b). *In vivo*, interaction between two human Rad52 molecules was demonstrated by the yeast two-hybrid assay. The result proposed that the self-interaction domain resides within the region of amino acid residues 65 – 165 of human Rad52 (Shen *et al.*, 1996b). In addition, *in vitro* binding assays with purified Rad52 protein further demonstrated that the region spanning amino acid residues 65 – 165 is necessary for Rad52-Rad52 interaction (Shen *et al.*, 1996b). Yeast Rad52 also forms ring-like structures *in vitro*. Using electron microscopy, Shinohara *et al*. demonstrated how yeast Rad52 protein covered FX174 ssDNA. They reported a “necklace” like structure with chains of Rad52 rings along the ssDNA. The ring-like structure of Rad52 has an outer diameter of 9 (±1) nm (Shinohara *et al.*, 1998). Electron microscopic analysis was also used by Van Dyck and co-workers to

show that human Rad52 protein self-associates to form ring-like structures with an approx. diameter of 10nm (Van *et al.*, 1998). The region responsible for Rad52 binding to itself has not been pointed out in yeast Rad52, but the region responsible for the self-association of human Rad52 shows 65% homology with the yeast sequence (Park *et al.*, 1996) making the N-terminal likely to be involved in yeast Rad52 self-association.

Electron microscopy was used to illustrate how wild type human Rad52 forms ring-shaped oligomers of various sizes (Ranatunga *et al.*, 2001). Work done by Stasiak *et al.* proposed that the human Rad52 ring is composed of seven monomers (Stasiak *et al.*, 2000). Ranatunga *et al*. reported that the size of the wild type Rad52 rings corresponded to 4 – 13 subunits with an average of 6. They also examined a *rad52* mutant expressing amino acid residues 1-192, corresponding to the part of the protein that has been crystallized. The mutant protein formed ring-like structures like the wild type protein (Ranatunga *et al.*, 2001), but the truncated Rad52-1-192Δ protein had particle sizes that corresponded to 4 – 15 subunits with an average of 10 subunits (Ranatunga *et al.*, 2001). Next, they constructed an N-terminal truncated *rad52* mutant expressing amino acid residues 218 – 418. Scanning transmission electron microscopy, gel filtration assays and dynamic light scattering suggested that the Rad52-Δ218-418 protein could self-associate and form oligomers. These results suggested the existence of an additional self-association domain in the C-terminus (Ranatunga *et al.*, 2001). The role of this domain could be to form higher order multimers of Rad52 protein.

It is still unknown whether the Rad52 ring is an active form of the protein. The ring formation may be a more general property among proteins involved in DNA metabolism, since it has been observed for a wide range of proteins like Rad51, Dmc1 and RecA. Van Dyck *et al*. have suggested that the Rad52 rings are active structures during single strand annealing of DNA mediated by human Rad52. Using electron microscopy, it was observed that Rad52 form rings on tailed DNA substrates, on annealed substrates and in larger complexes of DNA, which indicates a specific role of the ring structure (Van Dyck *et al.*, 2001). It has also been suggested that Rad52 binds DNA within the central channel of the heptamer like seen for human Dmc1 and SV40 large T antigen (Passy *et al.*, 1999), (Dean *et al.*, 1992). Van Dyck reported that, under the electron microscope, Rad52 appeared like beads on a string in the presence of ssDNA (Van *et al.*, 1998). Kagawa *et al.* and

Singleton *et al.* have suggested that Rad52 bind DNA in a groove of positive amino acid residues on the outside of the Rad52 ring-like structure (See figure 8) (Kagawa *et al.*, 2002), (Singleton *et al.*, 2002). In this thesis, a functional role of the Rad52 ring is proposed in chapter 2.

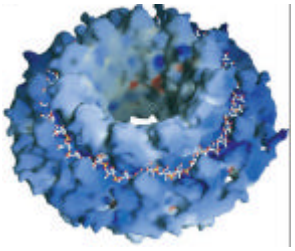


Figure 8. Predicted Rad52 ring-structure. The figure shows Rad52 molecules forming a ring-structure with the positive residues in red and the negative in blue. The large positively charged grove that runs across the protein could be a binding site for DNA (Singleton *et al.*, 2002).

1.5.5 Rad52 binds to other proteins

In addition to binding to itself, Rad52 also interacts physically with other proteins. Members of the *RAD52* epistasis group show many physical interactions, and Rad52 interacts specifically with Rad51, RPA and Rad59.

Rad52-Rad51 interaction

As mentioned above, Rad52 and Rad51 interact physically, both in yeast and human cells, and the binding capability has been attributed to the C-terminal of Rad52. However, the C-terminal has little homology between yeast and human (Milne and Weaver, 1993), (Shen *et al.*, 1995), (Shen *et al.*, 1996a), (Shinohara *et al.*, 1998) indicating that the Rad52 protein is functionally conserved in respect to protein-protein interactions. In yeast, Rad52 binds Rad51 through a few amino acid residues. Amino acid residues 409-412 in Rad52 are indispensable for Rad52-Rad51 interaction (Krejci *et al.*, 2002). Interaction of human Rad51 and Rad52 has been demonstrated both *in vivo* using the yeast two-hybrid system, and *in vitro* by immuno-precipitaton assays in insect cells expressing recombinant *RAD51* and *RAD52* (Milne and Weaver, 1993), (Hays *et al.*, 1995), (Mortensen *et al.*, 1996), (Shen *et al.*, 1996a). The interaction between Rad51 and Rad52 seems species-specific based on the observation that *rad51D* defects in *S. cerevisiae* not are complemented by *K. lactis* Rad51 or human Rad51 (Donovan *et al.*, 1994). Recent data underline the importance of Rad51-Rad52 interaction. Trace amounts of human Rad52 facilitates enhanced recruitment of Rad51 to the DNA targets (Navadgi *et al.*, 2005).

Krejci *et al.* have demonstrated that the interaction between Rad51 and Rad52 is physiologically important (Krejci *et al.*, 2002). A Rad52 mutant protein lacking five amino acid residues, Rad52-Δ409-412 is unable to interact with Rad51. The mutant protein was tested in biochemical assays. *In vitro*, Rad52-Δ409-412 behaved like wild type Rad52 in ssDNA annealing, DNA binding and RPA interaction assays. However, the mutated protein was devoid of mediator function in strand exchange assays demonstrating a specific physiological role for Rad52-Rad51 interaction. As a result, the mutant strain expressing the protein is defective for γ-ray damage repair, homologous recombination and for meiotic processes (Krejci *et al.*, 2002).

Rad52-RPA interaction

Both in yeast and human cells Rad52 interacts physically with the hetero-trimer RPA (Hays *et al.*, 1998), (Shinohara *et al.*, 1998), (Park *et al.*, 1996), and the nature of the Rad52-RPA interaction has been intensively explored in both organisms. Human RPA was found to associate through the middle part of the Rad52 protein both *in vivo* and *in vitro* (Park *et al.*, 1996), (Ranatunga *et al.*, 2001), (Jackson *et al.*, 2002). The RPA interacting region of human Rad52 has a limited amount of homology with yeast Rad52, thus it shows 7% identity to *S. cerevisiae* Rad52 sequence (Park, 1995), (figure 9).

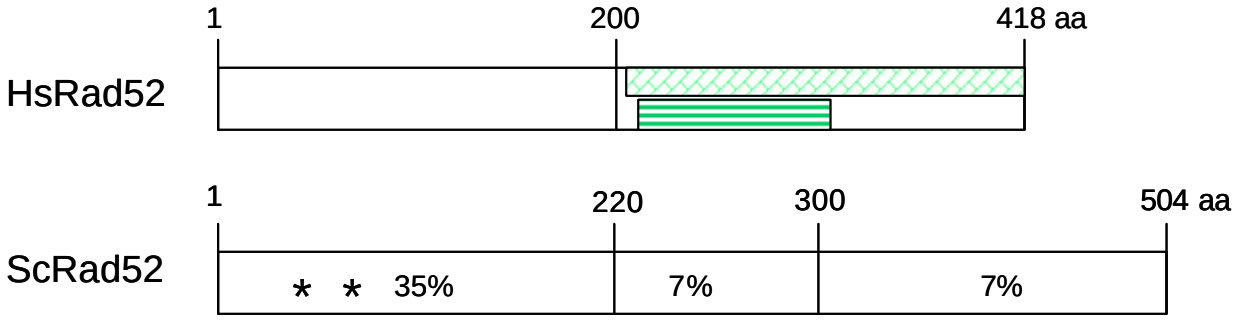


Figure 9. Conservation of amino acid sequences of human and yeast *RAD52*. The amino acid sequence homology between human and yeast Rad52 is shown in percentages on ScRad52 (Park, 1995). RPA interaction domain spanning aa 218-418 described by Ranatunga *et al*. is shown in squares (Ranatunga *et al.*, 2001). The domain described by Park *et al.* spans aa 221-280 and is shown in stribes

(Park *et al.*, 1996). ScRad52 interaction with RPA is disrupted by mutations G121E and G142D are shown in asterixes (Hays *et al.*, 1998).

In yeast, Hays *et al*. demonstrated an interaction of Rad52 with all three subunits of RPA. In addition, they proposed that the interaction between ScRad52 and subunits Rfa1 and Rfa3 is mediated through the N-terminal of Rad52. Two mutant *rad52* alleles, *rad52-34* (G121E) and *rad52-38* (G142D), were unable to interact specifically with Rfa1 and Rfa3, but still bound to the Rfa2 subunit (Hays *et al.*, 1998), (see figure 9). Yeast Rad52 co-precipitated with all three subunits of RPA. The amount of Rad52 that co-precipitated with RPA1 and RPA3 was significantly lower (0.4% and 1.6% respectively) than that for RPA2 (12%). This led the authors to suggest that the main interaction between yeast Rad52 and RPA occurs through the RPA2 subunit (Shinohara *et al.*, 1998). Human Rad52 interacts with the two largest subunits of RPA, but not with the smallest subunit. Like for ScRad52, the middle sized subunit appeared to have higher affinity to HsRad52 than the largest subunit (Park *et al.*, 1996).

In contrast to the work by Hays *et al.*, Ranatunga *et al.* pointed to the C-terminal part of HsRad52 as involved in RPA interaction. An enzyme-linked immuno-sorbent assay was used to show that human Rad52 expressing amino acid residues 218-418 interacts with RPA whereas Rad52 expressing amino acid residues 1 - 192 does not bind RPA (Ranatunga *et al.*, 2001), (figure 9). Human Rad52 has also been shown to interact with the 34 kDa subunit of RPA through a domain of Rad52 that is located between amino acid residues 221 – 280 (Park, 1995), (figure 9). It has not yet been established if the interaction of yeast Rad52 and RPA is mediated through the C-terminal like human Rad52 since no protein domain has been assigned for this function.

The association of human Rad52 and RPA has been shown to be functionally significant *in vivo*. *rad52* mutants lacking the RPA-interacting domain of HsRad52 were deficient in HR. Compared to cells expressing wild type *RAD52,* cells expressing the RPA-interaction deficient *rad52* strain did not induce homologous recombination (Park *et al.*, 1996).

Rad52-Rad59 interaction

A physical interaction between Rad52 and Rad59 was first reported in 2001. Davis and Symington used yeast two-hybrid and co-immunoprecipitation to demonstrate an

association between yeast Rad52 and Rad59 (Davis and Symington, 2001). Recently, a *rad52* mutant that is affected in its ability to interact with Rad59 has been identified. Accordingly, the mutant protein, Rad52-L89F, does not interact physically with Rad59-GST in a pull down assay (Cortes-Ledesma *et al.*, 2004).

1.5.6 Rad52, the mediator

Recent biochemical studies have shown that the Rad52 protein has at least two activities that are important in recombination. The role that Rad52 plays in annealing of complementary ssDNA in the recombination process has been described (Mortensen *et al.*, 1996), (Sugiyama *et al.*, 1998). A second role that Rad52 plays is as a mediator in the Rad51-dependent strand exchange reaction. Rad52 stimulates DNA strand exchange by Rad51 (New *et al.*, 1998) and overcomes the inhibitory effect of RPA that competes with Rad51 for binding ssDNA (Sung, 1997a). Rad52 helps Rad51 to displace RPA from ssDNA. Alone, Rad52 is not capable of this displacement, but it can form a complex with ssDNA-bound RPA, which recruits Rad51 to the ssDNA and RPA is displaced (Sugiyama and Kowalczykowski, 2002). Recently, new data has been published that suggests a novel role of Rad52 in strand exchange. Kumar and Gupta reported that human Rad52 also promotes strand exchange between two DNA molecules (Kumar and Gupta, 2004). Similar results were reported from Bi and co-workers, who also found Rad52 to promote the strand exchange reactions (Bi *et al.*, 2004).

1.6 The DSBR model – with proteins

The most important proteins in the DSBR pathway have now been introduced and the figure below summarizes the repair process in a simplified manner. The DSBR model is essentially the same model that was presented earlier (figure 2) with the most important proteins and their function added to it.

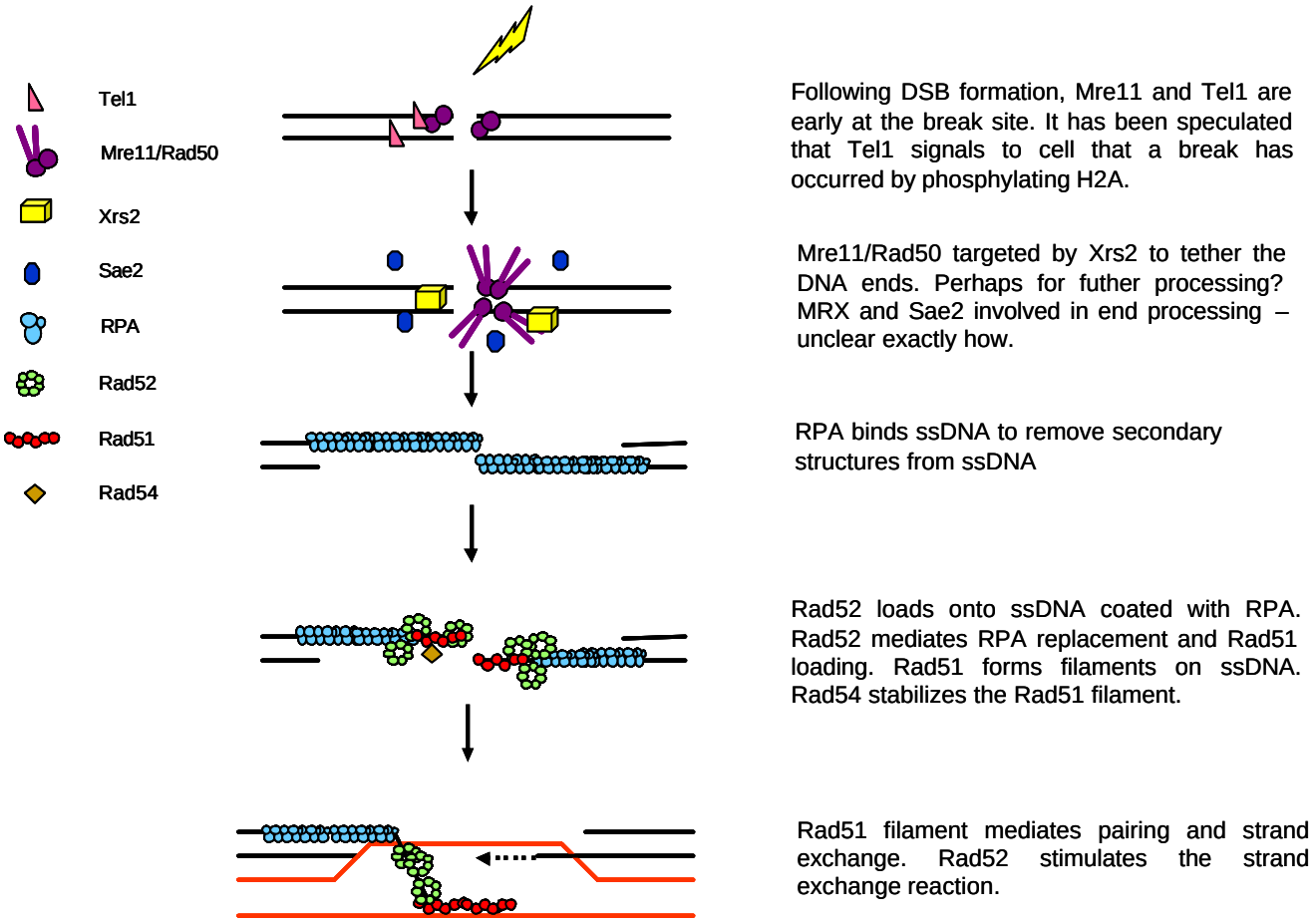


Figure 10 shows a simplified cartoon of the DNA DSB repair by the homologous recombination pathway. The figure is inspired from M. Lisby (Lisby *et al.*, 2004).

1.7 Nuclear localization of Rad52

As mentioned above, this thesis includes three research sub-projects. The first project to be introduced concerns how Rad52 is transported from the cytosol to the nucleus after translation. Though the knowledge on Rad52 is growing, it is unknown how this transport is mediated. Like other nuclear proteins, Rad52 needs to translocate from the cytosol to the nucleus to perform its normal cellular function. In the next section, the general mechanisms of nuclear transport of proteins in eukaryotes will be described together with a short introduction to computational methods used to predict sorting signals in protein sequences.

1.7.1 Transport factors

Most nuclear proteins larger than 40-60 kDa require active transport to enter the nucleus (Gorlich and Mattaj, 1996). Traffic of molecules across the nuclear membrane is tightly regulated and involves several transport factors and the mechanism has been reviewed intensely (Hicks and Raikhel, 1995), (Nakielny and Dreyfuss, 1999), (Macara, 2001), (Fahrenkrog and Aebi, 2003).

The nuclear import is mainly mediated by soluble proteins, the nuclear transport receptors that recognize the cargo and translocate it through the nuclear pore complex (NPC).

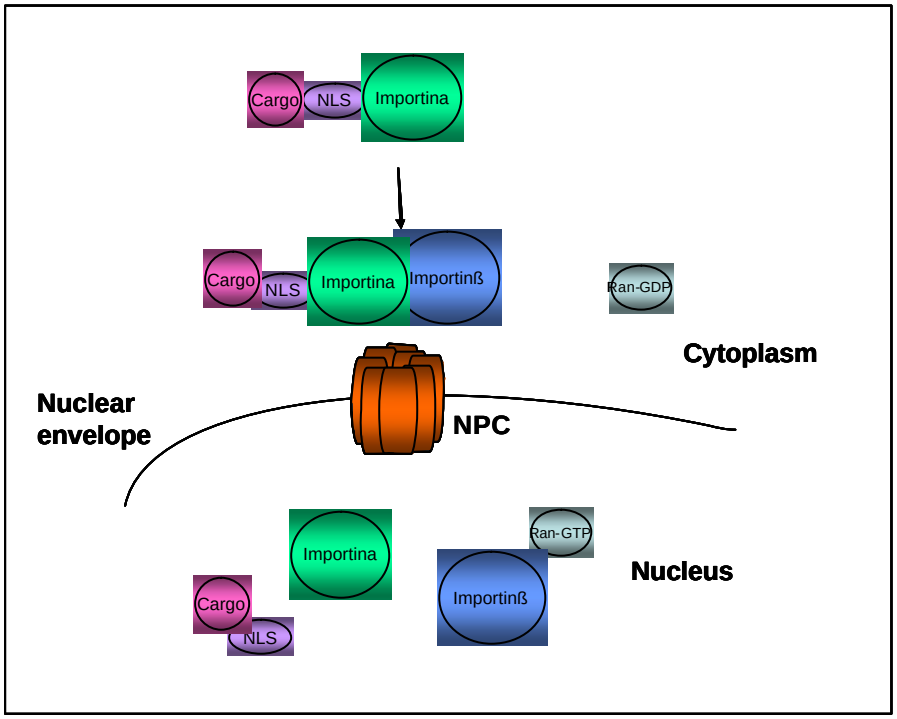


Figure 11. Nuclear import of a protein expressing a NLS. Importina binds the cargo protein via the NLS. Importinga then interacts with importinß that mediates transport through the NPC to the nucleus. In the cytosol the RanGTP concentrations are high and the cargo is released from the importin complex.

NLS sequences are recognized by importina of the importina-importinß heterodimer. Importina binds directly to NLS and mediates heterodimerization with importinß. The complex of the cargo protein and the importina-importinß heterodimer is docked at the NPC, and Importinß allows the translocation of the cargo through NPCs. The NPC span pores in the nuclear envelope and connect the nuclear and cytoplasmic compartments. The structure of the NPCs is conserved among species with its eightfold rotational symmetry around a cylindrical axis. In yeast, the NPC structure is comprised of approximately 30 different proteins, the nucleoporins. The NPC is composed of a nuclear and cytoplasmic ring, and eight spokes extend through the channel. It is not clear how the

cargo is propelled through the NPC, but it is generally believed that RanGTP is involved in this process.

RanGTP plays an important role in the transport of cargo across the nuclear membrane. The translocation through the NPC requires energy that is in part provided by Ran-catalyzed GTP hydrolysis. In addition to providing the required energy, the direction of the transport is also promoted by the Ran GTPase cycle. Ran is maintained as RanGTP in the nucleus by the guanine nucleotide exchange factor RCC1, and as RanGDP in the cytosol by the RanGTPase activating protein RanGAP. In the cytosol the concentration of RanGTP is low, and the import receptors bind to the cargo independent of RanGTP. However, when the cargo-RanGTP complex enters the nucleus where the concentration of RanGTP is high, the complex dissolves. Conversely, the export receptors bind their cargo when concentrations of RanGTP are high, like in the nucleus, and release the cargo in the cytosol when the RanGTP concentration decreases (Hicks and Raikhel, 1995), (Nakielny and Dreyfuss, 1999), (Macara, 2001), (Fahrenkrog and Aebi, 2003).

Specific transport in and out of the nucleus can be mediated by import and export signal sequences. Nuclear localization signals are generally characterized by one or more clusters of basic amino acid residues, but do not fit a tight consensus. The first NLS to be identified was in the simian virus 40 (SV40) large T antigen (P**KKKRK**V) (Kalderon *et al.*, 1984a), (Kalderon and Smith, 1984), (Kalderon *et al.*, 1984b). Classical NLSs are categorized as either monopartite, containing a single cluster of basic amino acid residues like the SV40 NLS, or bipartite, containing two clusters of basic amino acid residues separated by a linker of 10–12 un-conserved amino acids. The prototypical bipartite NLS is that of nucleoplasmin (**KR**PAATKKAGQA**KKKK**L) (Dingwall and Laskey, 1991).

Some proteins contain several NLS sequences. Presumably, more signals in a protein sequence increase the import of a protein to the nucleus because the rate of substrate interacting with importina increases (Boulikas, 1994). Other proteins contain so-called weak NLSs, which is NLS sequences that comprise stretches shorter than four basic residues or sequences where one or two basic residues have been replaced with histidine residues. These proteins are not transported to the nucleus as readily as proteins with

regular or multiple NLSs. Proteins with weak NLSs might require modifications like phosphorylation or require association with proteins with a strong NLS (Boulikas, 1997). A negative charge within the core NLS hexapeptide diminishes NLS function, but the NLS can be flanked by a negative amino acid without affecting the strength of the NLS. Moreover, NLSs can also be hidden inside the protein structure and be exposed by enzymatic cleavage of the protein, by dephosphorylation or by a conformational change in response to binding (Boulikas, 1994). Several nuclear proteins lack NLS, which suggest another mean of transport. The transport of these proteins can be dependent on interaction with proteins that express a functional NLS. There could be some advantages to the cell of having cytosolic association and subsequent nuclear import of a protein with an NLS and a protein without an NLS. One could envision two proteins that act together in the nucleus associate in the cytosol before entering the nucleus ensuring the right stoichiometric ratio of the two proteins (Boulikas, 1997).

1.7.2 Localization prediction

To date, several attempts to predict nuclear localization of proteins involved in DNA repair have been done. Boulikas has examined nuclear transport of DNA repair proteins in various organisms by identifying karyophilic clusters in these proteins (Boulikas, 1997). He manually searched the amino acid sequences for stretches containing four arginine, lysine, histidine or a stretch of four aspartic acid or glutamic acid residues within a hexapeptide (Boulikas, 1997) and defined a “typical NLS” as four lysine and arginine residues in a hexapeptide stretch (Boulikas, 1994), (see top box in table 1). Since an NLS often is located at a helix-turn-helix (Kalderon *et al.*, 1984a), (Kalderon *et al.*, 1984b), the a-helix breakers glycine and proline were also recorded. Boulikas concluded from his search, that ScRad52 is devoid of a putative NLS. However, he found that the homologue of Rad52 in *S. pombe* Rad22 contains a typical karyophilic cluster that might be responsible for the nuclear transport of Rad22 protein (Boulikas, 1997).

In addition to the manual approach, several algorithms and programs have been developed to predict subcellular localization of proteins. Such programs include ProtLock (Cedano *et al.*, 1997), a covariant discriminant algorithm (Chou K-C and Elrod D.W, 1999) and PSORT (Nakai and Kanehisa, 1992). We have successfully used the PSORT II algorithm (Nakai and Horton, 1999) to predict nuclear sorting signals in *S. cerevisiae* Rad52, and the program suggested the presence of two putative NLSs in the Rad52 sequence. In contrast, the ProtLock did not predict Rad52 to locate in the nucleus. The classification algorithm PSORT II is described in more detail in the appendix. PSORT uses the NLS of SV40 large T antigen to define an NLS, and two patterns are recognized as an NLS. The NLS of type 'pat4' is a residue pattern composed of four basic amino acids (K or R), like Boulikas definition or composed of three basic amino acids (K or R) and either H or P. The other type of NLS is referred to as 'pat7' and is a pattern starting with P followed, within three residues, by a basic stretch containing three K or R residues out of four (Hicks and Raikhel, 1995).

Table 1. NLS definitions. The definitions listed are defined by Boulikas manual NLS search and by the PSORT algorithm (Boulikas, 1997), (Nakai and Horton, 1999), (Hicks and Raikhel, 1995).

	Positions					
Boulikas	K/R	K/R	K/R	K/R		
	K/R	K/R	K/R		K/R	
	K/R	K/R		K/R	K/R	
	K/R		K/R	K/R	K/R	
		K/R	K/R	K/R	K/R	
	K/R/H	K/R/H	K/R/H	K/R/H		
'pat4'	K/R	K/R	K/R	K/R		
	K/R	K/R	K/R	H/P		
'pat7'	P	K/R	K/R	K/R		
	P		K/R	K/R	K/R	
	P			K/R	K/R	K/R

The algorithm in PSORT II is not as stringent as Boulikas definitions of possible NLSs (Boulikas, 1997) that did not predict any sorting signals in the Rad52 sequence. This raises a fundamental problem in prediction of signals based on amino acid sequences. It is important that the result from the algorithm provides a number of candidate sequences that can be investigated experimentally, and at the same time minimizes the number of false positives are kept low.

1.8 Rad52 repair center formation

Several repair proteins re-localize in response to DNA damage. This feature appears to be fundamental in DNA DSB repair, since it has been reported both in yeast, mouse, and human cells (Essers *et al.*, 2002), (Liu *et al.*, 1999), (Tan *et al.*, 1999). Hundreds to thousands of each repair protein accumulate at the break site to form bright dots, repair centers (Lisby *et al.*, 2003a), (Tan *et al.*, 1999). Still, it is unknown how these repair centers are formed in response to DNA damage and what the biological relevance is.

1.8.1 Methods to study DSBR

Much of what we know about the DNA DSBR process comes from numerous *in vivo* and *in vitro* assays that have become more and more sophisticated over the years. Biochemical assays provides important information about the function of the proteins outside of the living cell. However, the repair process *in vivo* is highly dynamic, which can be examined by fluorescent tagging of the proteins involved. Fusion proteins allow investigation of the ongoing repair process in living cells.

Functional proteins as well as DNA DSBs and repair sequences have been tagged *in vivo* with fluorescent probes. The break induction and repair in the cells has been examined by using fluorescence microscopy (Lisby *et al.*, 2003a). DNA DSB repair by HR has also been studied on synchronized cells by using inducible DNA DSBs based on site-specific endonucleases such as HO on a defined break site. Subsequent breaks and repair at the break site has been measured with southern blotting (Aylon and Kupiec, 2004). Furthermore, immunofluorescence experiments and chromatin immuno precipitation, ChIP, has been used to examine the repair process.

The formation of DNA DSB induced repair centers has been studied in detail in recent years *in vivo* and *in vitro*. Fluorescent probes make it possible to visualize DNA breaks, repair proteins and the focus formation directly at a single cell level (Essers *et al.*, 2002), (Lisby *et al.*, 2003a). Formation of repair centers is also visualized indirectly through immuno-staining with antibodies against the protein of interest (Essers *et al.*, 2002). In addition, the temporal order of recruitment to a focus has also been studied by ChIP

analysis (Wolner *et al.*, 2003). Additionally, a range of repair proteins from yeast and human cells have been purified and characterized biochemically (Erdile *et al.*, 1991), (Sung, 1997a), (Sung, 1997b), (Kantake *et al.*, 2003), (Solinger *et al.*, 2001), (Krejci *et al.*, 2003), (Mazin *et al.*, 2003).

Formation of repair centers in response to DNA damage has been a topic of great investigation and debate for the last few years and it is still an area of some dispute (Lisby *et al.*, 2003a), (Miyazaki *et al.*, 2004). Some research groups favor the thought that multiple DNA lesions are repaired concurrently in a repair factory where hundreds to thousands of repair proteins are assembled (Lisby *et al.*, 2003a). Other groups refuse this idea of a “recombinosome” and propose that DNA lesions are repaired one at a time in a one-lesion-one-repair-focus manner (Miyazaki *et al.*, 2004).

1.8.2 Rad52 repair centers in DSBR

As mentioned earlier, one aim of this thesis was to investigate how Rad52 is forming repair centers in response to DNA damage. *RAD52* alleles were tagged with YFP and subsequently the cellular distribution of the fusion proteins was visualized. There are several reasons why this method was used to study the active repair of DNA DSBs, and in the next section the advantages will be discussed.

.

A Rad52-YFP focus represents Rad52 engaged in the repair of a DNA DSB. Accordingly, Rad52-YFP is dispersed evenly in the nucleus but concentrates specifically in response to DNA damage. Rad52-YFP aggregates into repair centers after induction of DNA DSBs by γ-irradiation, meiosis or HO endonuclease (Lisby *et al.*, 2001). In addition, the concentration of Rad52-YFP is specifically located at a DNA lesion as shown by Lisby *et al*. Here, they visualize a significant co-localization of a single fluorescently labeled HO-inducible DNA DSB and Rad52 protein.

Rad52-YFP foci are formed spontaneously and turned over when the lesion has been repaired. Lisby *et al*. used time lapse microscopy to study the life time of Rad52-YFP repair centers. When there was a Rad52-YFP focus, they observed that the cell cycle halted until the Rad52 focus had disassembled (Lisby *et al.*, 2003b). This indicates that

disassembly of the Rad52-YFP repair center is necessary to de-activate the cell cycle checkpoint allowing the cell to proliferate.

The final argument in favor of using fluorescently tagged repair proteins to visualize DNA DSBR is the dynamic nature of the repair centers. The repair structures are flexible; proteins associate to and disassociate from the structure at different rates. Fluorescence loss in photobleaching (FLIP) and fluorescence recovery after photobleaching (FRAP) experiments demonstrate the mobility of the proteins in the repair complexes. Some proteins associated and disassociated rapidly to and from the complex whereas other proteins formed more stable structures (Essers *et al.*, 2002).

1.8.3 Why DSBR in repair centers?

One question to ask is whether the repair proteins build up into repair centers continuously or assemble in response to DNA damage? Dynamic structures allow flexibility in the composition of DNA repair complexes. Different components required for repair of different lesions and dynamic structures might allow exchange of different components in the complex (Essers *et al.*, 2002). Complexes assemble when required, which can benefit the cell and give rise to cross-talk between DNA repair pathways and coupling it to other DNA transaction such as replication (Essers *et al.*, 2002).

“On-the-spot” assembly of repair proteins can have some advantages to the cell since freely diffusing repair proteins in the nucleus provide the cell with a tool-box in close vicinity to the DNA and potential DNA lesions (Essers *et al.*, 2002). There is experimental data to back up the “on-the-spot” assembly hypothesis. *RAD52* group proteins do not co-immunoprecipitate in the absence of DNA damage indicating that the proteins are not pre-assembled in undamaged cells but only form in response to DNA damage (Tan *et al.*, 1999).

1.8.4 Chronology of the repair center build-up

After the cell has decided to repair a lesion by HR, DNA repair and checkpoint proteins are recruited to site of damage. This is performed in a highly organized manner that was

described recently by Lisby *et al*. (Lisby *et al.*, 2004). Fluorescently tagged repair and checkpoint proteins demonstrated that the order of recruitment was similar in spontaneous, endonuclease- and γ-irradiation-induced repair centers (Lisby *et al.*, 2004). A set of strains, disrupted in repair genes, were used to determine the repair center formation of fluorescently-tagged repair proteins. This genetic dissection of DNA repair center build-up showed that the recombination machinery is recruited to sites of DNA damage by Rad52 via its physical interaction with RPA. Below the interdependencies of checkpoint and repair proteins in recruitment to DNA damage will be described.

The repair and checkpoint proteins depend on the presence of RPA for repair center formation either directly, as Rad52, or through binding to Rad52 that again binds RPA. RPA on the other hand does not depend on the presence of any of the proteins investigated since it concentrates into repair centers in *rad51D*, *rad52D*, *rdh54D*, *rad54D*, *rad55D* and *rad59D* strains (Lisby *et al.*, 2004). This suggests that RPA is present at the DNA break before the other repair proteins. Accumulation of Rad52 at the break site might result in higher affinity for other proteins of the Rad52 group (Essers *et al.*, 2002). The central role of Rad52 in building repair centers can help to explain the severe defects in DNA DSB repair in *rad52* deficient cells (Lisby *et al.*, 2004).

In yeast, Rad51 is recruiting Rad55 to the site of damage by physical interaction. Rad55 is involved in the recruitment of Rad54 to the repair center. Rad54 is in part recruited by Rad51, but additional mechanisms are involved since Rad54 repair centers are not formed in a *rad55D* strain. Rad59 depends on Rad52 for recruitment, and fluorescently tagged Rad59 shows a cytosolic distribution in *rad52D* cells, which indicates that the nuclear transport of Rad59 is Rad52-dependent (Lisby *et al.*, 2004) and own observations.

However, the order of recruitment of Rad51 and Rad52 is a subject of controversy, which is outlined in figure 12. Fluorescence experiments by Lisby *et al*. proposed that Rad52 arrives at the break site before Rad51, and Rad51 is recruited to the repair focus via its interaction with Rad52. Accordingly, no Rad51 foci are formed in *rad52* deficient cells (Lisby *et al.*, 2004). This observation is in agreement with results by Sugawara *et al*. that also suggested that Rad52 is present at the break site prior to Rad51. Using ChIP

experiments at a specific DNA DSB they did not detect any significant immunoprecipitation of Rad51-associated DNA in a *rad52D* strain (Sugawara *et al.*, 2003).

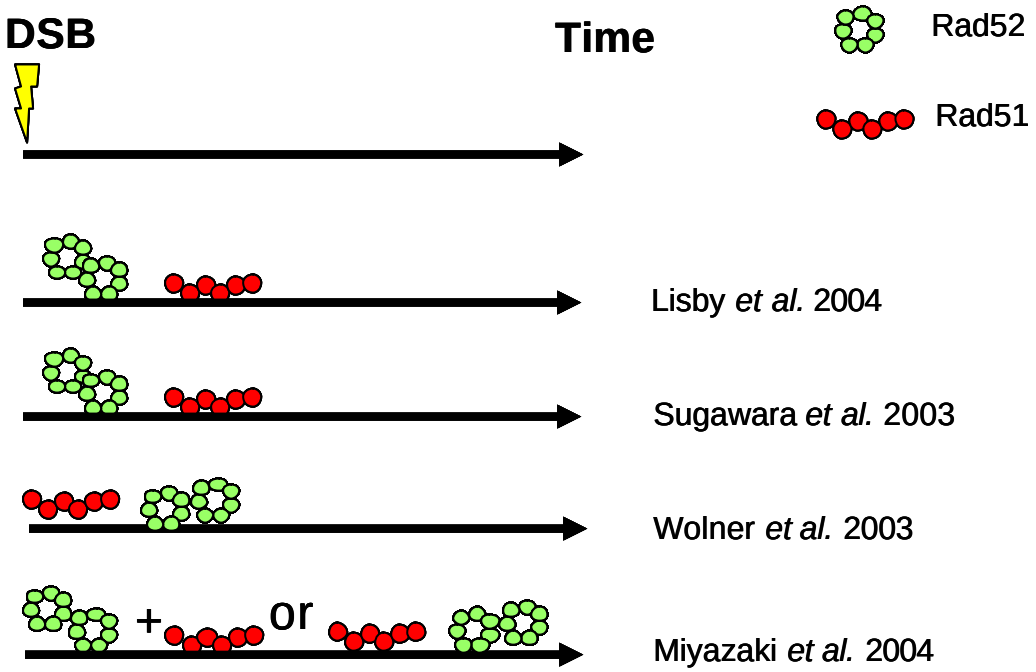


Figure 12. The order of Rad51 and Rad52 binding to the DNA break site.

Wolner *et al*. support an alternative order of recruitment following a unique DNA DSB in yeast (figure 12) (Wolner *et al.*, 2003). Their ChIP analysis showed Rad51 at a HO-induced DSB 20 minutes after induction and Rad52 present 40 minutes after the break was induced (Wolner *et al.*, 2003). The different results reported could be explained by the formation of an initial unstable binding of Rad52, which is not detected in a ChIP assay. After Rad51 binding, Rad52 arrives in levels that can be detected. This hypothesis was tested by monitoring the recruitment of the *RAD52* group proteins in *rad* mutant strains. Wolner *et al*. found Rad51 binding at the break site to be impaired in a *rad52D* background just like they found that Rad52 recruitment required Rad51 (Wolner *et al.*, 2003). This is in contrast to Lisby *et al*. that observed Rad52 foci in a *rad51D* strain but no Rad51 foci in a *rad52D* strain and interpreted this as Rad52 acting upstream of Rad51.

Furthermore, Miyazaki *et al*. have used nuclear spreads to investigate the repair center formation of Rad51 and Rad52 (Miyazaki *et al.*, 2004). They found Rad51 to assemble onto ssDNA simultaneously or slightly prior to the assembly of Rad52 and suggested that

Rad52 arrives at the break site before Rad51, but in levels below the threshold of detection. Although the levels are too low to be detected, the amount of Rad52 is still sufficient to promote the assembly of Rad51 onto the ssDNA. Another possible explanation is that the Rad52 present at the break is too unstable for detection, which makes it appear as if Rad51 is at the break site first. Alternatively, Rad52 and Rad51 could be recruited simultaneously (Miyazaki *et al.*, 2004).

Recent data from van Veelen *et al*. proposed that there exists a different DNA damage response in yeast and in mammalian cells (van Veelen *et al.*, 2005). They suggested that mammalian Rad51 repair centers are formed prior to Rad52 repair centers in response to DNA damage and that mammalian Rad52 is not required for ionizing radiation-induced repair center formation of Rad51 and Rad54 (van Veelen *et al.*, 2005). The question of order of recruitment of Rad51 and Rad52 is not definitely closed, but still needs to be addressed biochemically and genetically,

1.8.5 Dissociation of repair centers

When the DNA DSB has been repaired the repair center dissolves again. Lisby *et al*. have used time lapse microscopy to show how Rad52-YFP disassembles and allows the cell to proceed through the cell cycle (Lisby *et al.*, 2003b). The assembly and disassembly illustrates that it is indeed an ongoing process that is monitored by fluorescent microscopy. However, it is still unknown how the repair centers disassemble after the repair has been accomplished.

2. Identification of a nuclear localization signal in ScRad52

Rad52 is one of the key proteins in homologous recombination and disruption of the *RAD52* gene has severe phenotypical effects. Thus, *rad52D* strains are sensitive to ionizing-radiation, unable to sporulate and are genetically instable. Rad52 is a multi-domain protein and several functional regions in the protein have been mapped, including those involved in protein interaction and DNA binding. However, it is still unknown how Rad52 is transported to the nucleus following its translation since no putative NLS has been identified in Rad52. To investigate the possibility that a specific region in Rad52 is required for its nuclear transport a mutational strategy combined with a sequence analysis approach was employed. In this study, it was possible to identify an NLS in the middle region of Rad52, which is responsible for the nuclear transport of Rad52. Furthermore, the NLS was disrupted by one alanine substitution in the domain, which had no effect on the repair of MMS induced lesions suggesting that the role of the NLS in Rad52 is only to ensure the nuclear localization of Rad52.

2.1. Introduction

In the yeast *S. cerevisiae* mutations in *RAD52* show severe pleiotropic defects as Rad52 is required for all types of meiotic and mitotic HR including direct repeat recombination, break induced replication and telomerase independent extension of telomeres. The importance of Rad52 is further strengthened by the notion that Rad52 is conserved throughout the evolution and has been identified in numerous organisms like yeasts, fungi, fish, chicken, mouse and human (Adzuma *et al.*, 1984), (Sakuraba *et al.*, 2000), (van den *et al.*, 2001), (Takahashi and Dawid, 2005), (Bezzubova *et al.*, 1993), (Bendixen *et al.*, 1994), (Muris *et al.*, 1994), (Shen *et al.*, 1995), (Park, 1995).

Several biochemical properties of Rad52 are evolutionary conserved, both yeast and human Rad52 contains a DNA binding domain in the N-terminus that binds ss- and dsDNA (Mortensen *et al.*, 1996), (Shinohara *et al.*, 1998), (Van *et al.*, 1998), both promote annealing of complementary DNA strands (Mortensen *et al.*, 1996), (Sugiyama *et al.*, 1998), (Kagawa *et al.*, 2001), D-loop formation (Kagawa *et al.*, 2001), (Kumar and Gupta,

2004), (Arai *et al.*, 2005) and Rad51 catalysed strand exchange (Sung, 1997a), (New *et al.*, 1998), (Shinohara and Ogawa, 1998), (Kumar and Gupta, 2004), (Bi *et al.*, 2004).

Rad52 participates in several protein-protein interactions that are also evolutionary conserved. Both human and yeast Rad52 self-associate to form heptameric ring-like structures (Shen *et al.*, 1996b), (Stasiak *et al.*, 2000), (Ranatunga *et al.*, 2001), (Shinohara *et al.*, 1998) and both interact physically with Rad51 and RPA (Milne and Weaver, 1993), (Krejci *et al.*, 2002), (Shen *et al.*, 1996a), (Hays *et al.*, 1998), (Shinohara *et al.*, 1998). The position of the Rad51 interaction domain is located in the C-terminus of both *S. cerevisiae* Rad52 (ScRad52) and human Rad52 (HsRad52) despite that their sequences are quite dissimilar in this region (Milne and Weaver, 1993), (Krejci *et al.*, 2002), (Shen *et al.*, 1996a). The interaction domain responsible for RPA binding has been mapped to the middle section of HsRad52 (Park *et al.*, 1996), (Jackson *et al.*, 2002) whereas no region has been assigned to RPA interaction in ScRad52. However, point mutations that impair RPA binding have been identified in the N-terminus of ScRad52 indicating that this part of the protein may interact with RPA (Hays *et al.*, 1998). In addition, a Rad59 binding domain is present in the N-terminus of ScRad52 (Cortes-Ledesma *et al.*, 2004), (Davis and Symington, 2001). In HsRad52, a self-association domain that mediates higher order association of Rad52 rings has been proposed to reside in the C-terminus (Ranatunga *et al.*, 2001). Based on these analyses, a map of ScRad52 containing the different functions required for DNA repair can be drawn (figure 13). Above the drawing is shown the interactions demonstrated in ScRad52, below is shown the interactions known from HsRad52 (figure 13).

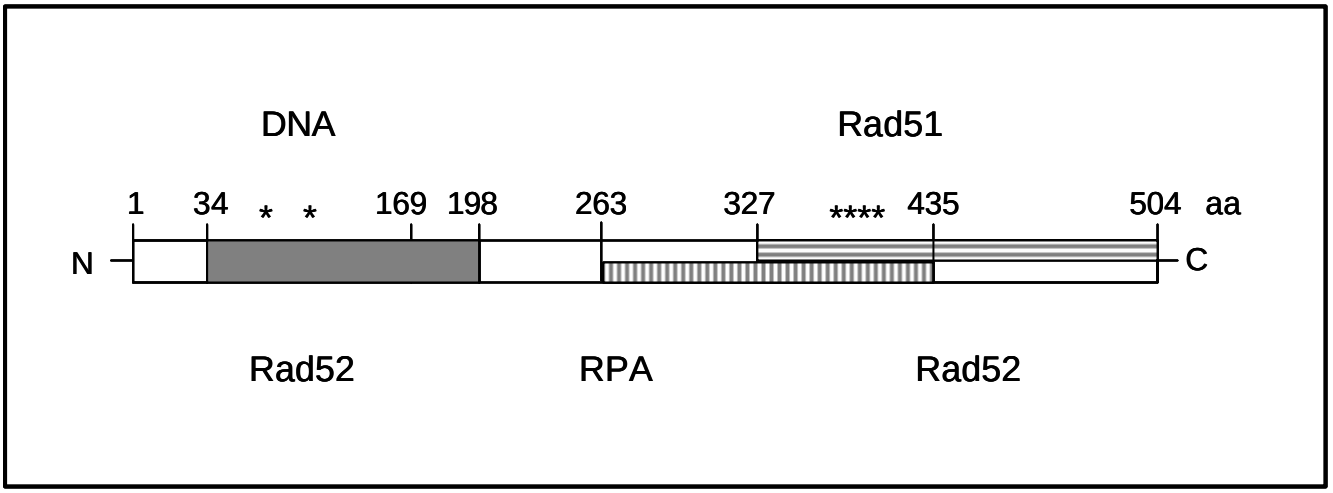


Figure 13. Functional map of *S. cerevisiae* Rad52

The outline shows the 504 aa of ScRad52. The dark shaded region spanning aa 34 – 198 corresponds to the evolutionary conserved region of ScRad52 (Mortensen *et al.*, 1996). DNA binding is mediated through the N-terminal, which is also responsible for HsRad52 self-association (Shinohara *et al.*, 1998), (Ranatunga *et al.*, 2001). A Rad51 interaction domain is located in the C-terminus. The horizontal stripes depict the Rad51 interaction domain identified by Milne and Weaver (Milne and Weaver, 1993) and in vertical stripes the domain Mortensen *et al.* found (Mortensen *et al.*, 1996). Amino acid substitutions G121E and G142D in ScRAD52 disrupting the interaction to Rfa1 (Hays *et al.*, 1998) are shown by asterisks in the N-terminus. The C-terminal asterisks mark the region including aa 409 -412 identified as responsible for Rad51 interaction by Krejci and co-workers (Krejci *et al.*, 2002). HsRad52 binds to itself through the N-terminus (Shen *et al.*, 1996b) as well as the C-terminus (Ranatunga *et al.*, 2001), and binds to RPA through the middle region (Park *et al.*, 1996), (Ranatunga *et al.*, 2001).

However, characterization of some of the important features in ScRad52 is still missing. For example, it is unknown what triggers transportation of Rad52 into the nucleus. The information needed for correct localization is generally found in the protein sequence, but no sequence in Rad52 has been shown to function as an NLS.

Most nuclear proteins larger than 40-60 kDa require active transport to enter the nucleus (Gorlich and Mattaj, 1996), (Macara, 2001). The size of Rad52 protein is 56 kDa (Shinohara *et al.*, 1998), placing it among proteins that require active transport. Furthermore, Rad52 may form ring structures in the cytosol resulting in large protein complexes too big for transport by simple diffusion. The active transport is facilitated by nuclear transport receptors, nucleoporins and importins that recognizes specific, in general

basic, import signals encoded in the sequence of the protein to be transported (Gorlich and Mattaj, 1996). The complex of a cargo protein and a nuclear transport receptor is then shuttled from the cytosol into the nucleus through the nuclear pore complex.

Since no putative NLS, has been identified in *S. cerevisiae* Rad52, it has been proposed that Rad52 is chaperoned into the nucleus via an interaction to another protein that has an NLS (Boulikas, 1997). In maps of mammalian Rad52, an NLS sequence has sometimes been included in its C-terminus. However, the functionality of this sequence has not been supported by experimental data.

Here, a biological functional Rad52-YFP fusion protein has been used to address cellular localization of Rad52. A *RAD52* deletion series has been constructed and their cellular distributions monitored by fluorescent microscopy, to delimit the region in Rad52 required for its nuclear localization. By combining the results from this analysis with a sequence analysis approach using the PSORT II algorithm (Nakai and Kanehisa, 1992), (Horton and Nakai, 1996) a region containing a possible candidate NLS was identified. Furthermore, alanine substitutions were introduced in this region to specifically identify amino acid residues required for nuclear transport. It was shown that a single NLS of type 'pat7' is essential for Rad52 transport into the nucleus. Proteins that are mutated in this sequence are biological functional. Accordingly, strains expressing these protein species survive chemically induced DNA DSBs if the mutant Rad52 species are extended by a functional NLS in the C-terminus. This strongly indicates that the sequence is important for nuclear localization only and not for Rad52 mediated repair functions.

2.2 Materials and Methods

2.2.1 Yeast strains and media

Yeast strains used in this study are displayed in table 2.

Table 2. Yeast strains used in this study

[a]Strain	Genotype	Source
UM94-5D	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 RAD52-YFP	U. Mortensen strain coll.
W2297-8C	Matalpha ADE2 his3-11,15 leu2-3,112 ura3-1 TRP1 lys? RAD52-YFP	U. Mortensen strain coll.
W3849-10C	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 TRP1 lys? RAD52-CFP	U. Mortensen strain coll.
UM94-9C	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52-?327-YFP	U. Mortensen strain coll.
UM94-2D	MATa ADE2 his3-11,15 leu2-3,112 1 ura3-1 trp1-1 LYS2 rad52-?307-YFP	This study
UM90-2C	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp 1-1 LYS2 rad52-?267-YFP	This study
UM91-4B	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp 1-1 LYS2 rad52-?237-YFP	This study
UMR188	MATa ade2-1 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52-?223-YFP	This study
UM69-1A	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52-?207-CFP	U. Mortensen strain coll.
UM56-7A	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52-?207-YFP	U. Mortensen strain coll.
UMR100	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52-?207-YFP-NLS	This study
UM101-15B	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52::HIS5	This study
UMR128	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52-R234A-YFP-NLS	This study
UMR131	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 RAD52-YFP-NLS	This study

[a]All strains are derivatives of W303-1A and W303-1B (Thomas, B. J. & Rothstein, R. 1989).
In addition to the genotype listed above all strains are RAD5

Table 2. Strains used in this study.

All media were prepared as described by Sherman (Sherman F, 1986) with minor modifications as the synthetic medium contained twice the amount of leucine (60 mg/L). All the strains used are isogenic to W303 (Thomas and Rothstein, 1989) except that they are RAD5 (Fan *et al.*, 1996), (Zou and Rothstein, 1997).

2.2.2 Plasmid construction

- *RAD52* expression vectors

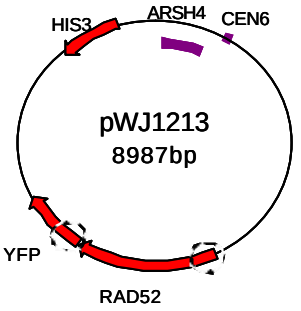


Figure 14. pWJ1213 is used as template for construction of KlRad52 and MmRad52 vectors. The C-terminal of ScRad52 and YFP are amplified and fused by PCR creating unique restriction cut sites for cloning.

A plasmid to express *K. lactis* (Kl) *RAD52* and *M. musculus* (Mm) *RAD52* was constructed from the plasmid pWJ1213 (kind gift from Michael Lisby). The sequence corresponding to the structural *RAD52* gene was removed and instead a fusion of the upstream region of *RAD52* (in frame with the *RAD52* promoter) and the start sequence of *YFP* was inserted. The insert was constructed by fusing two PCR fragments. The first fragment included the upstream region of *ScRAD52* up to the third start codon of *RAD52* and the second fragment included the start codon of *YFP*.

Fragments were generated by PCR using oligonucleotides AgeI_Rad52-fw: (5' – TGGATGGTACCGGTATCGAATGGCGTTTTTAAGCT) and Rad52_KasI_rv: (5'-aagtaacccgcggcgaacgaGGCGCCCAACAACACACCAAAGCCAC) for the first fragment, and BspEI_YFP_fw: (5' – tcgttcgccgcgggttacttTCCGGAATGAGTAAAGGAGAAGAACT) and BsrGI_YFP_rv: (5' - CTGATTGCTGTACA-TAACCTTCGGGCATGGCACTC) for the second. A unique cut site for *Kas*I was introduced in the first fragment, *Bsp*EI in the latter. A complementary tag sequence of 20 nucleotides at the end of each fragment facilitated fusion of the two PCR products in a subsequent PCR reaction (tags shown in small letters). PCR was performed using the Expand^{TM} high fidelity system according to the supplier's instructions (Roche Diagnostics). The PCR construct was inserted into pWJ1213 vector using *Age*I and *Bsr*GI to construct vector pIPL1.

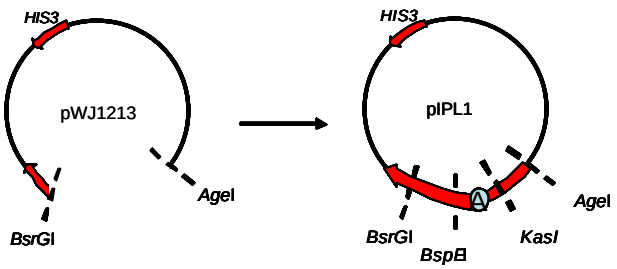


Figure 15. Constructing pIPL1. Digestion with *Age*I and *BsrG*I leaves the vector pWJ1213 with the upstream region of *RAD52* and the 3' end of *YFP*. Upstream of *RAD52* until third start codon and start of *YFP* is amplified and fused with PCR and re-inserted in vector to make pIPL1. Unique restriction cut sites for cloning *Kas*I and *BspE*I are introduced in pIPL1 for subsequent cloning.

RAD52 from *Kluveromyces lactis* and *Mus musculus* were amplified from *K. lactis* CBS2359 (originating from the Centraal Bureau voor Schimmelcultures, Delft, The Netherlands) (Kiers *et al.*, 1998) and pWJ669 (a kind gift from Rodney Rothstein) (Bendixen *et al.*, 1994) respectively. The oligonucleotides used were Kl_Rad52_KasI: (5' – CGTCCGTAGGCGCCAATGGAGGATACAGGAAGTGGC) and Kl_Rad52_rv: (5' – CCGTGTGCTCCGGATGACACATTTCTTCTGTTTG) for the *KlRAD52* fragment, and Mu_Rad52_KasI: (5' – CGTCCGTAGGCGCCAATGGCTGGGCCTGAAGAAGC) and Mu_Rad52_rv: (5' – CCGTGTGCCCCGGGCCCGGATGGATCTAGTTTCCTTTTC) for the *MmRAD52* fragment. Both genes were constructed to be out of frame with the upstream region of *ScRAD52* and lacking stop codons to ensure fusion to YFP. The *KlRAD52* gene was inserted into pIPL1 using *Kas*I and *Bsp*EI, and the *M. musculus* gene was inserted using *Kas*I and *Xma*I.

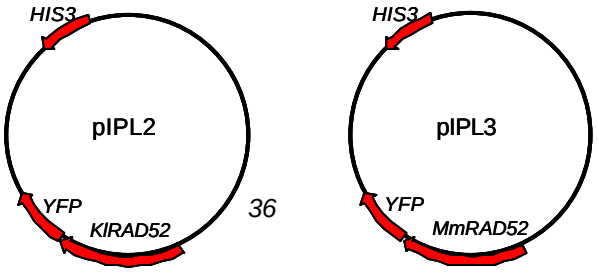


Figure 16. Construction of pIPL2 and pIPL3. *RAD52* from *K. lactis* and *M. musculus* are cloned into pIPL1 using *Kas*I/*BspE*I and *Kas*I/*Xma*I, respectively. The corresponding constructs, pIPL2 and pIPL3 harbors *Kl*RAD52-YFP and *Mm*RAD52-YFP, respectively.

The new vectors, pIPL2 and pIPL3, harboring *RAD52-YFP* from either *K. lactis* or *M. musculus* were transformed into a *RAD52* strain (UM94-5D) and a *rad52D* strain (UM101-15B) and nuclear localization was visualized by fluorescent microscopy.

- pWJ1213-*rad52-K233A-R234A-R235A-YFP* and pWJ1213-*rad52-R234A-YFP*

Two distinct NLS 2 mutants were made by substituting selected basic amino acids of the NLS 2 sequence (PNKRRQL) with alanine residues. In one mutant, all of the three basic amino acids (K233, R234 and R235) were substituted with alanine residues. The other mutant contained single amino acid substitutions of residue R234.

To construct a PCR fragment, where the basic amino acids of the NLS 2, K233, R234 and R235, has been substituted with alanine residues, two PCR fragments were generated with the primers N-term-seq2 (5' GAGTGTGGCAACGCCAGCAGTG -3') and Rad52-233-35-3ala-adap-Rev. (5' CTTTAGTCAATTGTGCGGCCGCATTCGGGTATTGTTGTTGTTCTTG -3') in one reaction and with the primers Rad52-233-35-3ala-adap-Fw (5'-GCGGCCGCACAATTGACTAAAGTTACAAATACCAATCCCG-3') and Rad52_down (5'- AATGAACCTAAGGATTCCGC-3') in another reaction. For both reactions the plasmid PRS413-RAD52 was used as template. The two fragments were fused in a third PCR reaction with the primers N-term-seq2 and Rad52_down. Fragments containing the mutation R234A, were made in the same way, except the mutating primers were Rad52-R234-F (5'- aaggcgcgcCAATTGACTAAAGTTACAAATACC -3') and Rad52-R234-R (5'- ctttagtcaattggcgcgccttATTCGGGTATTGTTGTTGTTCTTG -3'). PCR fragments were inserted into pWJ1213 with *Age*I and *Sph*I.

- pWJ1213-*rad52-D207-237-YFP* and pWJ1213-*rad52-D207-237-YFP-NLS*

An internal deletion of amino acid residues 208 to 237 in RAD52 was constructed by two rounds of PCR to obtain Rad52-Δ207-237-YFP. The deletion was incorporated in the

upstream fragment with oligos Nterm_Seq2: (5' – GAGTGTGGCAACGCCAGCAGTG), and 207-238-rv: (5' – ttggtatttgtaactttagtCCTAAACAAATTGTTTTCGT) and in the downstream fragment with oligos 238-fw: (5' – ACTAAAGTTACAAATACCAA) and Rad52-down: (5' – AATGAACCTAAGGATTCCGC) (adaptamer sequences are shown in small letters). The two PCR fragments were fused by their adaptamer sequences in a subsequent PCR reaction using oligos Nterm_Seq2 and Rad52-down. The RAD52 fragment was inserted into pWJ1213 by AgeI and SphI resulting in a fusion of Rad52-Δ207-237 to YFP.

A PCR fragment containing *rad52-D207-237-YFP-NLS* was constructed by PCR with oligos Nterm_Seq2 and YFP-NLS-Stop-XhoI: (5' – CCGTGTGCCTCGAGTCACTCGACTTTCCGCTTTTTCTTCGGAGATGCTTTGTATAGTTCATCCATGC) using pWJ1213-r*ad52-D207-237-YFP* as template. The fragment was inserted into vector pWJ1213-*rad52-D207-237-YFP* using *Age*I and *Xho*I.

- pWJ1213-*rad52-169D-YFP* and pWJ1213-*rad52-169D-YFP-NLS*

The construction of pWJ1213-*rad52-169D-YFP* and pWJ1213-*rad52-169D-YFP-NLS* is described in materials and methods in chapter 3.

2.2.3 Construction of genome integrated *RAD52-YFP* fusion mutants

Genomic integrated *RAD52* mutants were constructed and fused to *YFP* using the cloning-free PCR-based allele replacement method previously described (Erdeniz *et al.*, 1997;Lisby *et al.*, 2001). Rad52 was tagged with YFP at the C-terminus, mutated or truncated using appropriate primers. *URA3* from *K. lactis* was used as a counter-selective marker in the integration process. For the truncation mutants the downstream region of *RAD52* was amplified with oligos RAD52Tdown: (5' – ggatgaactatacaaataaCCCGCTTCCTGGCCGAAAC) and Rad52_down. 2/3 of *K. lactis URA3* and *YFP* was amplified from pWJ1165 (Lisby *et al.*, 2001) with oligos YFP_Tdown: (5' – TCCCCGCGGTTATTTGTATAGTTCATCC) and 5'int: (5' – CTTGACGTTCGTTCGACTGATGAGC). The *RAD52* downstream fragment was fused to the *URA3-YFP* fragment with oligos 5'int and RAD52-down. *YFP* and the other 2/3 of *K. lactis URA3* was amplified from pWJ1164 (Lisby *et al.*, 2001) with oligos YFP_startF: (5' –

ATGAGTAAAGGAGAAGAAC) and 3'int: (5' – GAGCAATGAACCCAATAACGAAATC). This fragment was fused to the upstream *RAD52* fragments encoding differing numbers of Rad52 amino acid residues.

Oligos for construction of mutant *rad52-D223-YFP* encoding the first 223 amino acid residues of Rad52 were Rad52start-F: (5' – ATGGCGTTTTTAAGCTATTT) and Rad52_223_reverse_A: (5' – AGTTCTTCTCCTTTACTCATTTCATGCAAAGTATTCGTTC). Oligos for *rad52-D237-YFP* encoding the first 237 amino acid residues of Rad52 were Rad52start-F and Rad52_237_reverse_A: (5' – agttcttctcctttactcatCAATTGCCTTCTTTTATTCG). Primers for *rad52-D267-YFP* encoding 267 amino acid residues of Rad52 were Rad52start-F and Rad52_267_reverse_A: (5' – agttcttctcctttactcatTGCGGCCATCATTGGGGTAC). Shown in small letters is an adaptamer sequence complementary to the first 20 nucleotides of *YFP* including a stop codon after *YFP*. The fragment containing the 5' 2/3 of *K. lactis URA3*, *YFP* and downstream *RAD52* is co-transformed with fragments containing the 3' 2/3 of *K. lactis URA3*, *YFP* and *RAD52* of different sizes.

The substrates were integrated at the endogenous *RAD52* locus via sequence homology and transformants expressing functional *K. lactis URA3* were selected on media lacking Uracil. Rare recombination between the two *YFP* repeats removed one *YFP* repeat and the *K. lactis* marker, which could be selected for on 5'-fluoroorotic acid (5-FOA) media. All *RAD52* mutants were confirmed by sequencing (MWG Biotech AG).

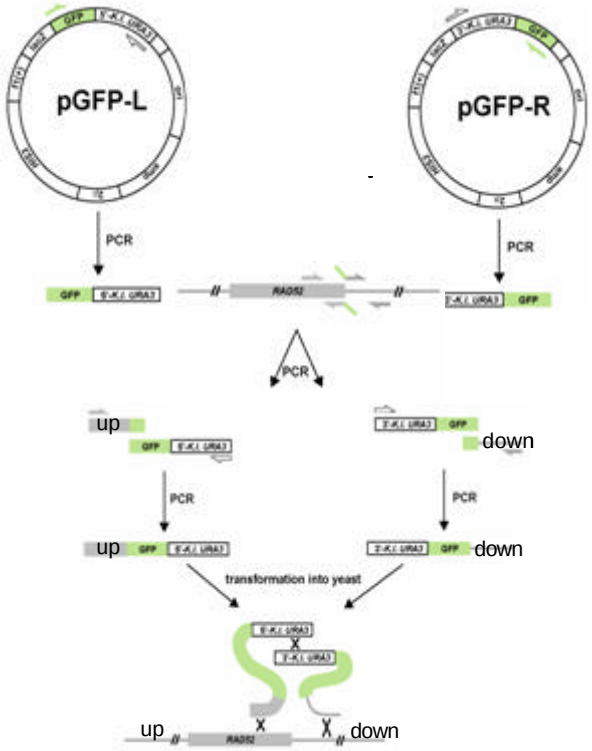


Figure 17. The PCR based gene allele replacement method.

Genomic integration of *rad52-R234A-YFP* and *rad52-R234-YFP-NLS*

A strain encoding Rad52-R234A-YFP was constructed by transforming the *rad52-D207-YFP* strain (UM56-7A) with a bipartite substrate consisting of a fragment encoding *rad52-R234A-YFP* fused to 2/3 upstream part of the *K. lactis URA3* marker and a fragment encoding the 2/3 downstream part of *K. lactis URA3* marker fused to *YFP*. A PCR fragment containing *rad52-R234A* was amplified with the oligos Rad52start-F: (5' – ATGGCGTTTTTAAGCTATTT – 3') and Urtag1: (5' – gttcttctcctttactcatCCCAGTAGGCTTGCGTG - 3') using pWJ1213-*rad52-R234A-YFP* as a template (the lower-case characters of the Urtag1 sequence indicate the adaptamer complementary to the first 19 nucleotides of YFP). The adaptamer sequence was used to fuse the *rad52-R234A* fragment to the *YFP* fragment fused to the 2/3 downstream part of *URA3*. Fragments containing YFP-5' 2/3 URA3 and YFP - 3' 2/3 URA3 were constructed as described in the text above. The strain was named UMR128. A strain encoding *rad52-R234A-YFP-NLS* was constructed by transforming the *RAD52-YFP-NLS* strain (UMR130) with the same bipartite substrate used to construct UMR128.

2.2.4 MMS assay

To assess sensitivity to the alkylating agent methyl methanesulfate (MMS) (M4016 from Sigma), the mutagenized plasmids were transformed into a *rad52D* strain (UM101-15B). Cells were grown overnight at 30 °C in selective media (SC-His), washed with sterile water and resuspended in an appropriate volume. Subsequently, six 10-fold dilutions were made of cell suspensions containing 10^8 cells per ml and 5µl of each dilution was spotted on SC-His plates containing 0 %, 0.0025 % and 0.005 % MMS. Plates were incubated at 30 °C for 2 days before examination (Prakash and Prakash, 1977a).

2.2.5 Microscopy

In all experiments, cells were grown at 23°C (to allow the YFP and CFP chromophores to form efficiently) in 3 ml overnight cultures to an OD of 0.3 at 600 nm. Fluorophores used in this study were yellow- and blue-shifted enhanced variants of the *GFP* gene (Ormo *et al.*, 1996;Heim and Tsien, 1996). In the experiment with integrated *RAD52* mutations cells

were grown in SC medium, in the experiment investigating *RAD52* mutants expressed from a plasmid cells were grown in SC-His medium to select for the plasmid.

For all epifluorescence microscopy, cells were spun down and immobilized on a glass slide by mixing with a 37°C solution of 1.2% (wt/vol) low melting agarose (NuSieve 3:1 from FMC) containing the appropriate medium. For MMS experiments, the overnight cultures were spun down and resuspended in 0.5% MMS, and then incubated for 15 min. Next, cultures were washed twice with water and resuspended in fresh SC or SC-His medium as appropriate. The cells were then incubated for 30 min to allow for foci to form before imaging. Selected strains were made rho^0 (mitochondrial DNA negative) before staining to eliminate any signal from mitochondrial DNA. DNA was stained for visualization by adding 10ug/ml DAPI to the culture for 12 min prior to imaging.

The filters used to visualize Rad52-GFP (excitation 480 nm; emission 535 nm) and 49,6-diamidino-2-phenylindole (DAPI; excitation 365 nm, emission 450 nm) were from Omega Optical (Brattleboro, VT). Images were acquired by using a cooled CCD camera (Star-1 from Photometrics, Tucson, AZ) mounted on a Zeiss Axioplan II microscope with a Plan-Apochromat 1003, 1.4 numerical aperture (NA) objective lens. The illumination source was a 100-W mercury arc lamp. Integration time for acquisition of the fluorescent images was 200 ms. Images were acquired in the IP LAB software (Scanalytics, Billerica, MA) and ADOBE PHOTOSHOP (Adobe Systems, Mountain View, CA).

2.2.6 Construction of diploid heterologous *RAD52* strains with *CFP* and *YFP* fusions

rad52-D207-CFP (UM69-1A) was crossed with UM91-9D, UM93-17D, UM94-6C and W2297-8C to generate diploids *rad52-D207-CFP/rad52-D237-YFP*, *rad52-D207-CFP/rad52-D267-YFP*, *rad52-D207-CFP/ rad52-D307-YFP* and *rad52-D207-CFP/ RAD52-YFP*. Diploid cells were grown ON in SC media and the subcellular localization of the proteins examined by using fluorescence microscopy.

2.3 Results

2.3.1 *K. lactis* Rad52 locates in the nucleus in *S. cerevisiae*.

Rad52 is a multi-domain protein that interacts with DNA, itself and several other protein partners. Only the N-terminus, which contains the domains involved in binding to Rad59, binding to DNA, and self-association is highly conserved. In contrast, the Rad51 binding domain in HsRad52 and ScRad52 and the RPA binding domain in HsRad52 are positioned in regions that show only little sequence identity between Rad52 species identified from different organisms.

To determine if the domain responsible for nuclear localization of Rad52 is sufficiently evolutionarily conserved to retain interspecies functionality, *CEN6* based vectors encoding *K. lactis* and *M. musculus* Rad52 species tagged with YFP were constructed. In both plasmids, *RAD52* expression was controlled by the *S. cerevisiae RAD52* promoter (figure 8A). The two plasmids along with a *S. cerevisiae* Rad52-YFP control plasmid, were individually transformed into wild-type and *rad52D S. cerevisiae* strains and the transformants were inspected by fluorescent microscopy to determine the cellular localization of the Rad52 species. Both KlRad52-YFP and MmRad52-YFP were expressed in *S. cerevisiae* in sufficient levels to detect the cellular localization of the fusion proteins. KlRad52-YFP was mainly located in the nucleus of *rad52D* strains Figure 18 shows how KlRad52-YFP locate in the nucleus of *S. cerevisiae rad52D* cells. Pictures of DAPI stained cells show that the KlRad52-YFP and the nucleus co-localize. In agreement with this result, it has previously been demonstrated that expression of KlRad52-YFP in *S. cerevisiae* partly suppresses the severe phenotype of *rad52D* strains (Milne and Weaver, 1993). In contrast, most MmRad52-YFP remains in the cytosol of *rad52D* strains and only a very faint YFP signal is picked up in the nucleus (figure 18).

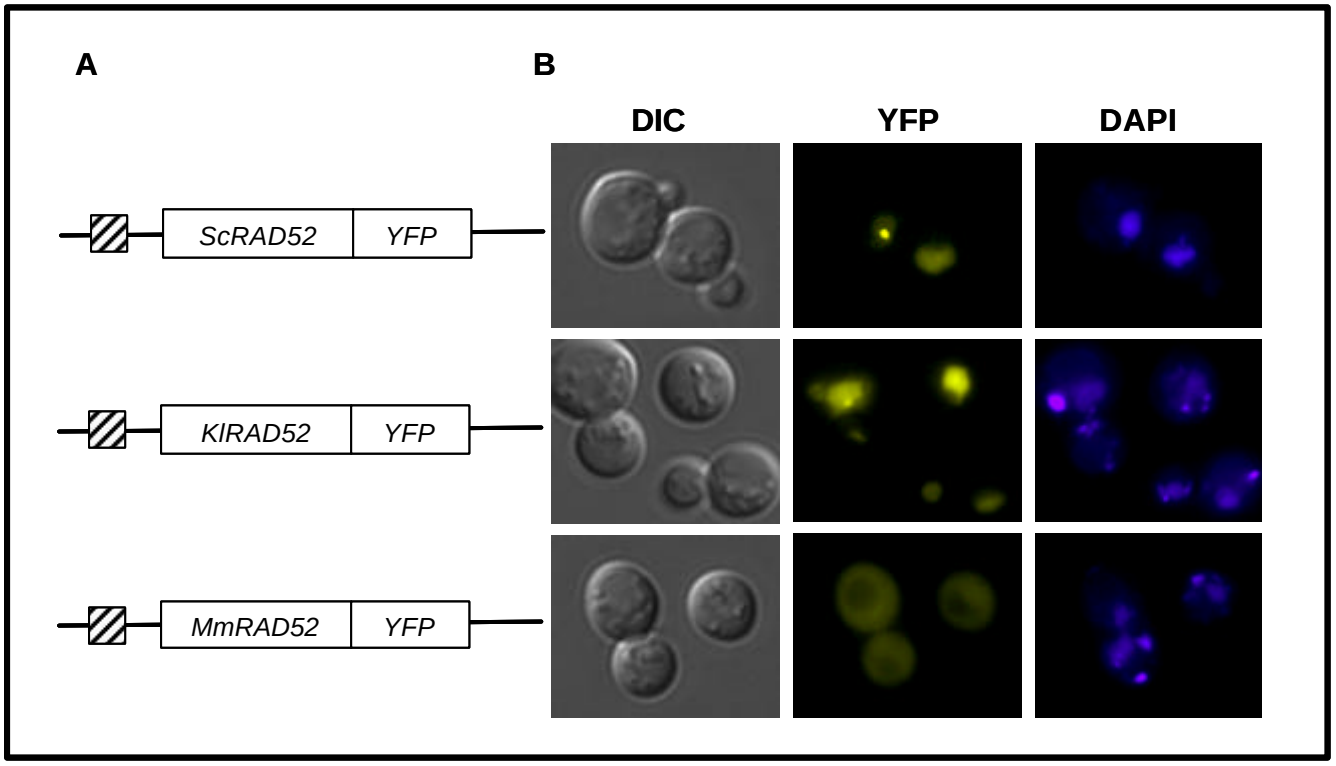


Figure 18. *KlRAD52* and *MmRAD52* expressed in *S. cerevisiae*. A. schematic representation of the *RAD52* constructs shows the *RAD52-YFP* genes from different species expressed from a CEN6 plasmid harboring the *S. cerevisiae* promoter (indicated by the hatched box). *RAD52* fragment from *Mus musculus* was cloned into vector pIPL1 by *Kas*I and *Bsp*EI, for the *KlRAD52 Kas*I and *Xma*I were used. **B**. Fluorescent microscopy of *rad52D* cells expressing Rad52-YFP fusion proteins from *S. cerevisiae*, *K.lactis* and *M. musculus*. Pictures shown are pseudocoloured images; Rad52-YFP in yellow and DAPI stained DNA in blue. Localization of Rad52-YFP and derivatives here from was not affected by the absence of mitochondria or by DAPI staining (data not shown).

These results suggest that a common mechanism ensures nuclear localization of ScRad52 and KlRad52, but not MmRad52. If so, sequence motifs important for nuclear localization of ScRad52 are likely to be present in KlRad52, but not in MmRad52. Co-expression of ScRad52 and MmRad52 did not result in localization of MmRad52 to the nucleus (data not shown).

2.3.2 A sorting signal is located in the middle-part of Rad52.

Manual sequence searches as well as bioinformatics have previous been unsuccessful to predict NLS sequences in ScRad52 (Boulikas, 1997), (Cedano *et al.*, 1997). However,

analysis using the PSORT II software and *k* nearest neighbor classifier (Nakai and Kanehisa, 1992), (Horton and Nakai, 1996) predicts two putative NLSs in the N-terminal of the ScRad52 sequence spanning aa residues 148 - 151 and 231 – 237, respectively (Figure 19). The predicted NLSs are RRKP (aa 148 – 151) of type 'pat4' and PNKRRQL (aa 231 – 237) of type 'pat7' (Hicks and Raikhel, 1995). The NLSs are referred to as NLS 1 and NLS 2.

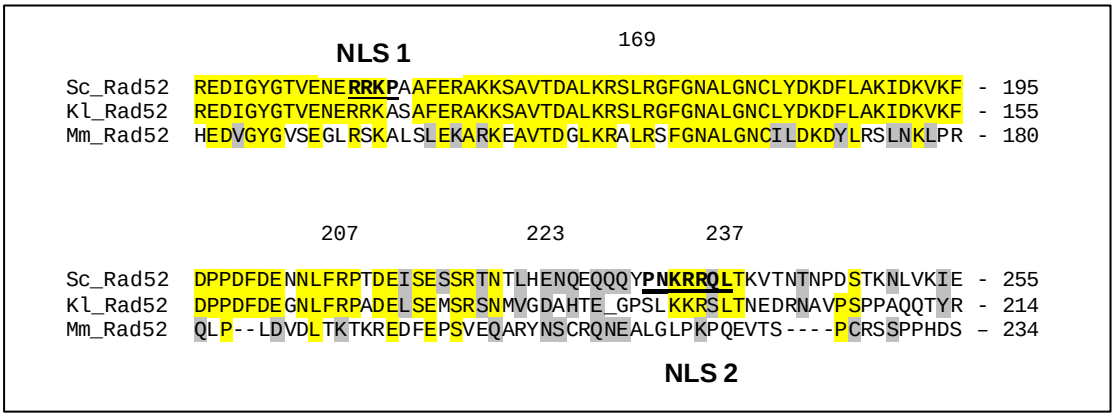


Figure 19. Predicted NLSs in *S. cerevisiae* Rad52. The figure shows the two predicted NLSs in the *S. cerevisiae* Rad52 sequence located at aa 148 -151 (NLS 1) and aa 231 – 237 (NLS 2), respectively. Identical amino acid residues are depicted in yellow and similar residues in grey. The bold, underlined stretches indicate the predicted NLSs.

The experimental approach to map the region of Rad52 required for nuclear localization was to construct a series of *S. cerevisiae* strains where the endogenous *RAD52* locus has been modified to express five different Rad52 C-terminal truncations all fused to YFP (figure 20A).

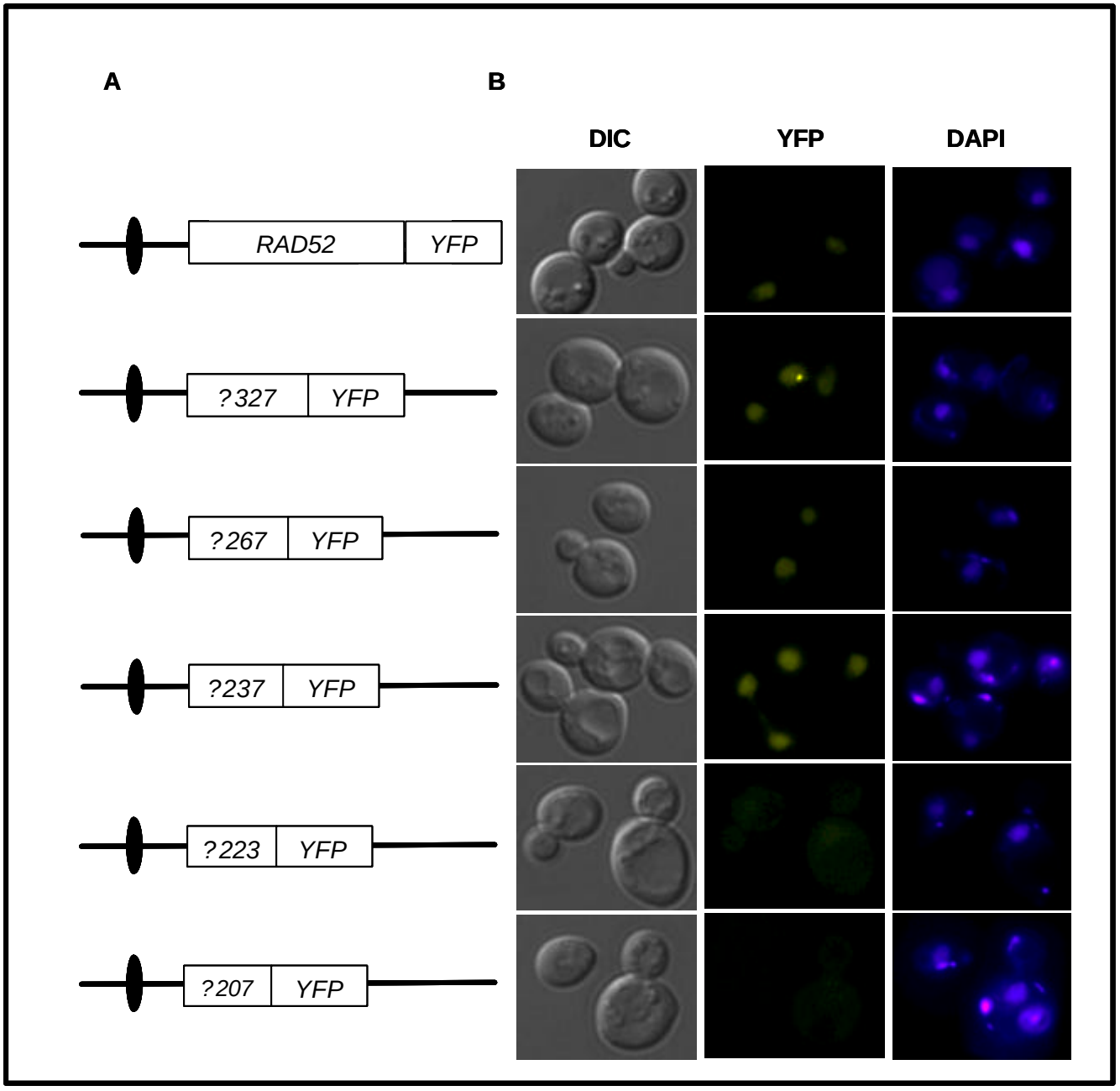


Figure 20. Diagram of integrated YFP-tagged *S. cerevisae RAD52* alleles. **A**. The 504 amino acid coding *RAD52* gene and carboxy-terminal truncations encoding 327, 267, 237, 223, and 207 amino acid residues. **B**. The nuclear localization of the YFP-fused *RAD52* alleles.

All strains were subjected to fluorescent microscopy to determine the cellular location of the fusion proteins. Not surprisingly, Rad52-YFP localized in the cell nucleus (figure 20). Of the truncations, Rad52-*D*327 was also expected to sort into the nucleus. This is because the strong MMS sensitivity of a *rad52-D327* strain previously has been shown to

be fully suppressed by overexpression of *RAD51* (Milne and Weaver, 1993), (Boundy-Mills and Livingston, 1993). In contrast, *rad52* mutant strains expressing larger C-terminal truncations were only partially suppressed by Rad51 overexpression (Asleson *et al.*, 1999). This phenotype could be due to elimination of Rad52 function or elimination of efficient nuclear localization (or both). In agreement with this, it was found that Rad52-Δ327-YFP (the smallest C-terminal truncation) localizes in the nucleus, whereas Rad52-Δ207-YFP (the largest C-terminal truncation) mostly localizes in the cytosol (figure 20B). Rad52-Δ207-YFP is also present in the cell at low concentrations, which makes it difficult to determine if there is a slight up concentration of the mutant protein in the nucleus. The cellular location of the three remaining C-terminal deletion truncations was analyzed to further delimit the region important for nuclear transport. Of these, Rad52-Δ237-YFP and Rad52-Δ267-YFP localize in the nucleus and Rad52-Δ223-YFP localizes in the cytosol (figure 20B). Again, the mis-sorting fusion protein Rad52-Δ223 is expressed at low levels compared to the nuclear localized protein species. However, it is still possible to see that the Rad52-Δ223-YFP is mostly in the cytosol, but also in small concentrations in the nucleus.

Together these results indicate that a region involved in nuclear sorting is located between amino acid residues 207-237. To support this conclusion, a CEN6 based vector was constructed, which allowed expression of Rad52 species, Rad52-Δ207-237-YFP, where this internal region was deleted.

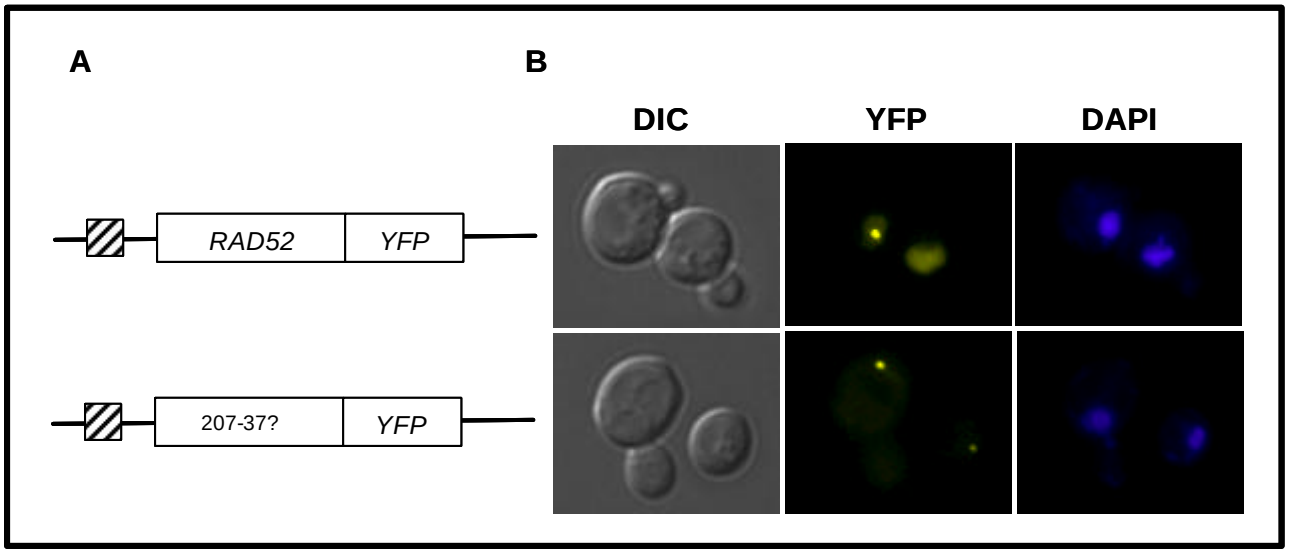


Figure 21. Rad52-Δ207-237-YFP locates in the cytosol. A. schematic representation of *RAD52-YFP* and *rad52-Δ207-237-YFP* constructs **B**. Fluorescent microscopy of *rad52Δ* cells expressing Rad52-YFP and

Rad52-Δ207-237-YFP. Pictures shown are pseudocoloured images; Rad52-YFP in yellow and DAPI stained DNA in blue.

As expected, this Rad52 species localizes in the cytosol (figure 21). However, the fusion protein does also form inclusion bodies in the cytosol. The mis-sorting of Rad52-Δ207-237-YFP suggests that NLS 2, one of the NLS sequences predicted PNKRRQL (aa 231 – 237) is responsible for the nuclear sorting of Rad52. Interestingly, KlRad52, but not MmRad52, also contains a similar sequence, PSLKKR, (figure 19) perhaps explaining why KlRad52 and not MmRad52, is sorted to the nucleus of *S. cerevisiae* in a *rad52D* strain.

2.3.3 Amino acids 231 - 237 is an NLS that ensures nuclear transport of Rad52

Three basic aa residues (K233-R234-R235) constitute the core of the predicted NLS 2 identified in ScRad52. To verify that these amino acid residues function as an NLS, two additional plasmids were constructed. The first construct allowed expression of the triple mutant Rad52-K233A-R234A-R235A-YFP and the second the expression of the single mutant Rad52-R234A-YFP. Both plasmids were transformed into *rad52D* strains and the resulting transformants investigated by fluorescent microscopy (figure 22).

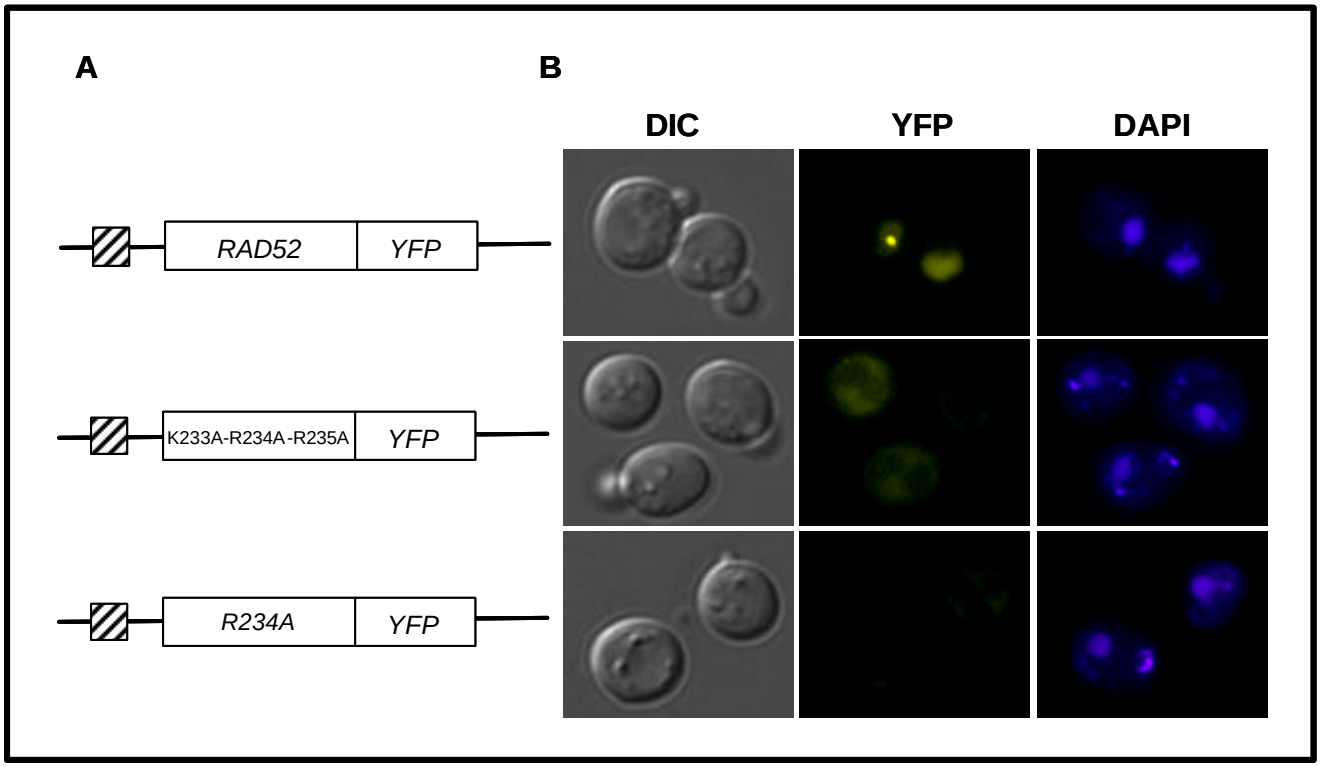


Figure 22. Rad52-K233A-R234A-R235A-YFP and Rad52-R234A-YFP locate in the cytosol. A. schematic representation of *RAD52-YFP, rad52-K233A-R234A-R235A-YFP* and *rad52-R234A-YFP* constructs. The hatched box represents the *RAD52* promoter. **B**. Fluorescent microscopy of *rad52D* cells expressing Rad52-YFP, Rad52-K233A-R234A-R235A-YFP and Rad52-R234A-YFP. Pictures shown are pseudocoloured images; Rad52-YFP in yellow and DAPI stained DNA in blue.

The figure shows the wild type Rad52-YFP located in the nucleus. Neither Rad52-K233A-R234A-R235A-YFP nor Rad52-R234A-YFP was sorted correctly to the nucleus as they mainly were located in the cytosol (figure 22). Both mutant proteins are present in the cell at lower concentrations than wild type Rad52-YFP, form inclusion bodies and are unstable. However, the localization result was confirmed by integrating the R234A mutation into a strain where the endogenous *RAD52* gene was extended by *YFP* (data not shown). The alanine substitutions in the predicted NLS (NLS 2, aa 231-237) suggest that these amino acid residues are indeed indispensable for the nuclear localization of Rad52.

2.3.4 Amino acids 148 - 151 are not sufficient to ensure nuclear transport of Rad52

Next, the putative NLS at aa residues 148 -151 (NLS 1) was investigated to see if it influenced nuclear sorting of Rad52. Since Rad52-Δ207-YFP expresses NLS 1, but does not locate in the nucleus it does not appear to be sufficient for nuclear localization. However, it can be because this particular NLS 1 is not responsible for the nuclear transport or that Rad52-Δ207-YFP mis-fold or that the fluorescent tag masks the transport of the truncated protein. However, when Rad52-Δ207-YFP was tagged in the C-terminal with NLS from SV40 virus, it lead to translocation of Rad52-Δ207-YFP-NLS to the nucleus (figure 25) indicating a partly folded Rad52-Δ207-YFP as it is recognized by the nuclear transport receptors and transported to the nucleus.
It is therefore concluded that the single “pat7” NLS, NLS 2, spanning aa residues 231-237 is responsible for sorting Rad52 to the nucleus.

2.3.5 Rad52 nuclear transport domain has no effect on DNA DSBR

Next, it was investigated whether the role of the NLS motif is solely in nuclear transport and not in the DNA DSBR function of Rad52. Accordingly, Rad52-Δ207-237-YFP was

tagged C-terminally with the well characterized NLS sequence from SV40 virus (Kalderon *et al.*, 1984a), (Kalderon and Smith, 1984), (Kalderon *et al.*, 1984b) to mediate its nuclear transport by a non-Rad52 sequence. Inspection of the resulting strain by fluorescent microscopy showed that the fusion protein indeed concentrated in the nucleus (figure 23) and concentrated to form bright fluorescent foci. Figure 23 shows wild type Rad52-YFP localized in the nucleus as well as Rad52-Δ207-237-YFP-NLS and the mis-sorting Rad52-Δ207-237-YFP localized in the cytosol.

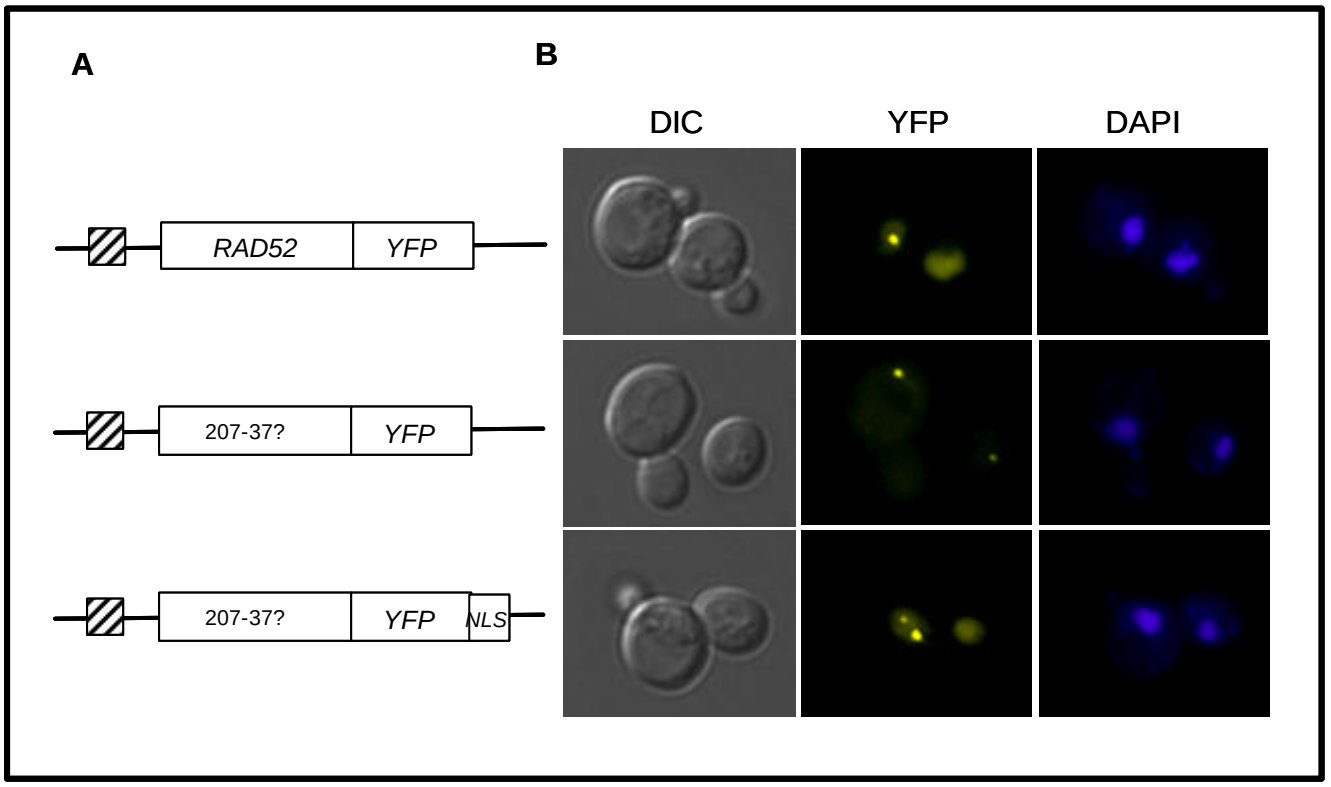


Figure 23. Rad52-*D*207-237-YFP-NLS locates in the nucleus. A. schematic representation of *RAD52-YFP, rad52-D207-237-YFP* and *rad52-D207-237-YFP-NLS* constructs. **B**. Fluorescent microscopy of *rad52D* cells expressing Rad52-YFP, Rad52-Δ207-237-YFP and Rad52-Δ207-237-YFP-NLS. Pictures shown are pseudocoloured images; Rad52-YFP in yellow and DAPI stained DNA in blue.

The ability of Rad52-Δ207-237-YFP-NLS to perform Rad52 functions during repair of MMS induced damage was determined. A *rad52D* strain was transformed with plasmids expressing either *RAD52-YFP*, *rad52-D207-237-YFP*, *rad52-D207-237-YFP-NLS* or an empty vector and spotted on MMS containing plates.

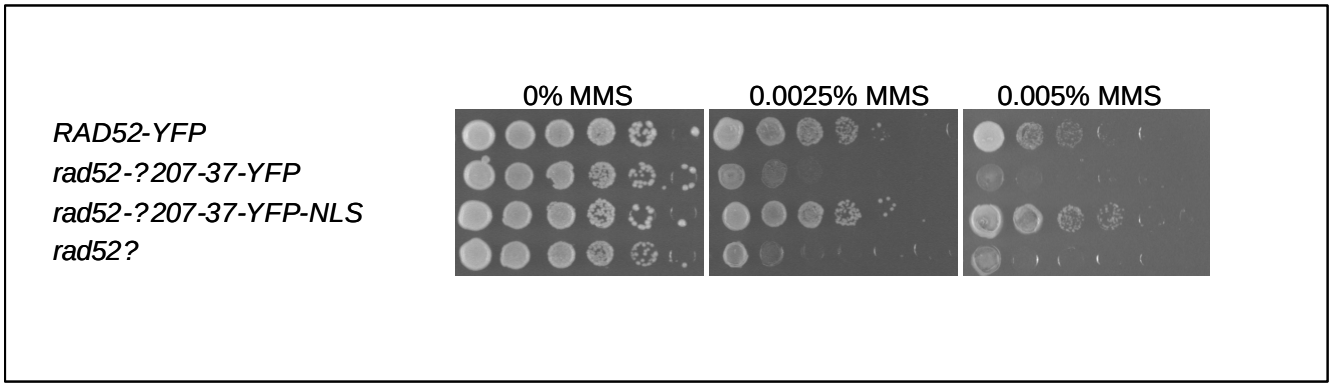


Figure 24. Survival assay shows region spanning aa 207 – 238 of Rad52 to be dispensable for repair of MMS induced lesions. 10^8 cells/ml were diluted in a 10-fold series and spotted onto SC-His plates containing 0, 0.0025 and 0.005 % MMS. Pictures were captured after two days incubation.

The survival assay showed that NLS 2 is dispensable for repair of MMS-induced DNA damage. At the MMS dose used, the viability of *rad52D* and *rad52-D207-237-YFP* strains was strongly reduced whereas wild-type *RAD52* and *rad52-D207-237-YFP-NLS* strains were almost unaffected (figure 24). This result shows that the entire NLS containing region can be eliminated without affecting repair of MMS damage significantly suggesting that the role of the NLS signal in Rad52 only is to ensure the nuclear localization of Rad52**.**

2.3.6 Co-expression of sorting and mis-sorting Rad52 species results in nuclear localization of both proteins

The identified NLS 2, PNKRRQL, is considered a weak NLS because it comprises a stretch shorter than four basic residues (Boulikas, 1997). However, the presence of more weak NLSs in a protein may increase the overall signal and allow a more readily nuclear transport of the protein (Boulikas, 1994).

If ScRad52 contains a single weak NLS this raises the possibility that Rad52 by self-association in the cytoplasm forms heptameric rings that will have seven weak NLS signal sequences, which together ensures efficient nuclear transport. If so, one would expect that

a wild-type Rad52 species may assist nuclear entry of a Rad52 species without a functional NLS, as they should form chimeric Rad52 ring-structures containing both species in the cytosol (figure 25). Figure 25B illustrates how an NLS-containing Rad52 molecule interacts physically with a Rad52 molecule without an NLS and forms a chimeric ring-structure.

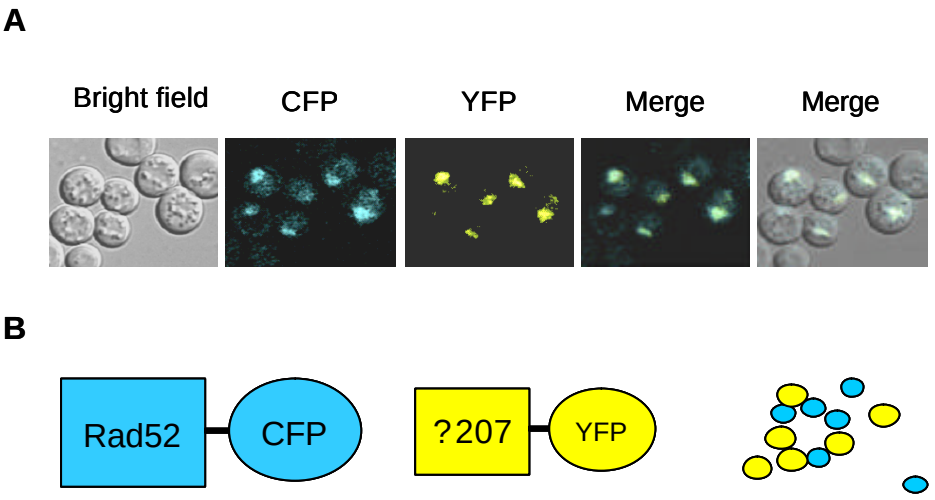


Figure 25. Co-expression of Rad52 species with and without NLS.

A. When Rad52-*D*207-YFP and wild type Rad52-CFP are co-expressed, Rad52-Δ207-YFP locates in the nucleus together with Rad52-CFP. **B**. Model suggesting that Rad52-Δ207-YFP and Rad52-CFP form a chimeric ring-structure facilitating nuclear transport of both protein species.

To test this possibility, diploid strains co-expressing wild type Rad52-CFP and YFP-tagged Rad52 mutant proteins were constructed. The NLS-defective species Rad52-Δ207-YFP and Rad52-Δ223-YFP co-localized with Rad52-CFP in the nucleus (figure 25 – data for Rad52-Δ223-YFP and the rest of the Rad52 truncations are found in appendix). This result indicates that there is a physical association between the Rad52 molecules in the cytosol prior to the nuclear sorting of the proteins (figure 25B). Co-expression of Rad52-Δ207-YFP with Rad59 did not result in a translocation of the Rad52 mutant indicating a specific Rad52-Rad52 interaction (data not shown).

2.3.7 Monomeric Rad52 species unable to sort to the nucleus

Lastly it was determined whether multimerization is necessary for efficient nuclear transport of Rad52. It is hypothesized that a ring-structures is established in the cytosol prior to nuclear transport. Consequently, a Rad52 species, which lacks the N-terminal self-

association domain, but contains the region where the proposed NLS is located would not form a ring-like structure and mis-sort to the cytosol. Such an allele, Rad52-169Δ-YFP expressing amino acids 169 - 504 fused to YFP locates in the cytosol. However, when tagged C-terminally with NLS from SV40 virus it locates in the nucleus. This suggests that Rad52 is transported to the nucleus as a multimer, since the monomeric Rad52-169Δ-YFP mis-sorts and locates in the cytosol.

2.4 Discussion

In this study, a mutagenesis of *RAD52* was performed and combined with a sequence analysis approach to identify a region responsible for nuclear transport of Rad52. Using these methods it was possible to identify a "pat7" type NLS in the middle region of the Rad52 sequence that is necessary for efficient nuclear transport of Rad52 and the middle region was dispensable for repair of MMS induced DNA DSBs. However, the results also indicate that the NLS is not sufficient to mediate efficient nuclear transport of Rad52 species lacking the N-terminal self-association domain. Thus, it was proposed that Rad52 proteins form a ring-like structure in the cytosol prior to the nuclear transport.

The sequence analysis pointed at two regions in Rad52 as potential NLS sequences, NLS 1 at aa 148 -151 and NLS 2 at aa 231 – 237 and they were analyzed to determine their involvement in Rad52 nuclear transport. This analysis involved results from fluorescent microscopy of the series of fluorescently tagged Rad52 truncation and deletion mutants and showed that NLS 2 is important for nuclear transport of Rad52 since alanine substitutions in NLS 2 disrupted nuclear transport of the mutant protein without changing its ability to repair MMS induced damage.

The other putative NLS, NLS 1, was not sufficient to ensure successful nuclear transport of Rad52 since Rad52-Δ207-YFP does not sort to the nucleus even though the protein sequence includes NLS 1. In addition, the Rad52 species *rad52-D210* that also expresses NLS 1 fails to sort efficiently to the nucleus. But the mutant protein is not exclusively excluded from the nucleus, since Asleson and co-workers showed truncated protein species to partly complement a *rad52D* strain (Asleson *et al.*, 1999). *rad52-D210* showed 2-fold more resistance to MMS than *rad52D* and demonstrated that at least a fraction of the truncated protein Rad52-Δ210 is present in the nucleus. Other results however,

indicated that NLS 1 does not play a central role in Rad52 nuclear transport. A previous screen of a large collection of mutants included mutations in the region of NLS 1. One of these mutants, *rad52-K150A* was only mildly sensitive to γ-ray, exhibited direct repeat recombination rates like wild type, but low recombination rates (Mortensen *et al.*, 2002). This indicated that the mutant protein Rad52-K150A sorted to the nucleus, where it is present at concentrations sufficient for γ-ray induced damage. However, the separation-of-function phenotype might be due to different parts of Rad52 being involved in DNA repair and in recombination. The results by Mortensen *et al*. indicated that NLS 1 is not solely responsible for the nuclear transport of Rad52. If it was the case, it would be expected that disruption of the signal would lead to defects in Rad52 function.

Likewise, mutant proteins expressing NLS 2, Rad52-169Δ-YFP are unable to localize to the nucleus. An NLS sequence is apparently not sufficient to ensure transport to the nucleus and here it is suggested, that Rad52 also forms a ring-like structure in the cytosol before it enters the nucleus. The observations from the mis-sorting fusion protein Rad52-Δ207-YFP support this hypothesis. Rad52-Δ207-YFP is unable to sort correctly to the nucleus, but when it is co-expressed with wild type Rad52 it relocates to the nucleus together with the wild type protein. It could be because the two protein species form a ring-like structure in the cytosol that mediates the nuclear transport. If Rad52 associates in a ring of seven molecules it provides a total of seven weak NLSs, which can facilitate a more readily nuclear transport than if the proteins were to be transported as monomers.

It is not sufficient to express an NLS to ensure nuclear localization. Likewise, the ability of a Rad52 species to form ring-structures is not in itself enough to facilitate nuclear transport. In addition, the protein species also has to express NLS 2 or interact with a protein species that expresses NLS 2. To illustrate this, a *RAD52* mutant that expresses both NLS 1 and NLS 2, but lacks the N-terminal Rad52 self-association domain was constructed. This protein species did not locate in the nucleus, but remained in the cytosol, probably because the mutant protein failed to associate with other Rad52 mutant molecules and therefore remained monomeric. This observation was supported by Lumir Krejci. He showed that Rad52-M, which includes NLS 1 in the sequence and Rad52-MC that include NLS 1 and NLS 2 to be monomeric. It suggests that the existence of NLS

sequences in the Rad52 sequence is not sufficient to ensure nuclear transport. In addition, the protein also has to be engaged in a ring-like structure prior to transport to ensure proper nuclear transport. The reason for mis-sorting of these protein species could also be mis-folding, but this is unlikely since they all sorted correctly to the nucleus when tagged C-terminally with the strong SV40 NLS.

We suggest that multimerization of Rad52 in the cytosol is necessary for its nuclear transport and propose a model (see figure 26) where Rad52 monomers assemble in the cytosol prior to translocation to the nucleus. Thus, Rad52 can be transformed to the nucleus by two mechanisms. Rad52 can either sort to the nucleus by expressing NLS 2 and associate into a ring prior to the nuclear transport. The green Rad52 molecules on the figure below illustrate a ring-like structure of Rad52 proteins expressing both NLS 2 and the N-terminal Rad52 interaction domain. Rad52 species can also associate with an NLS-containing species and thereby piggyback its way to the nucleus. Figure 26 illustrates how Rad52 proteins expressing NLS 2 as well as the N-terminal Rad52 interaction domain (in green) interact physically with mis-sorting Rad52 mutant proteins without NLS 2, but with the N-terminal interaction domain (in blue). Together this chimera is sorted to the nucleus.

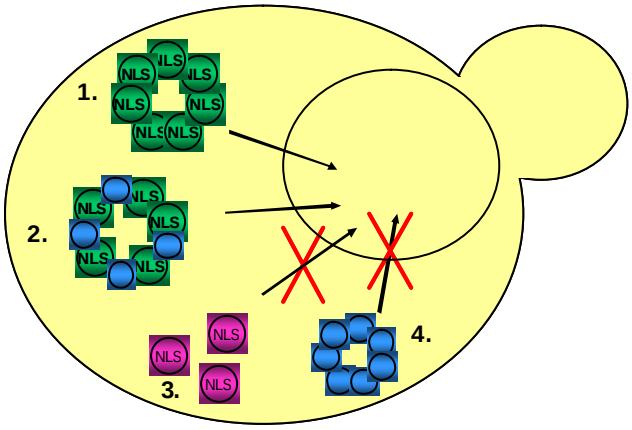


Figure 26. Model for nuclear transport of Rad52. Rad52 can be transported in two ways; **1)** if the Rad52 species encodes an NLS in the protein itself (green) and is able to self-associate or **2)** by piggy-bagging on another Rad52 protein that has an NLS. The mis-sorting Rad52 is shown in blue. Rad52 protein species fail to sort to the nucleus if; **3)** the protein encodes an NLS, but lacks the Rad52 N-terminal interaction domain (pink) or **4)** if the Rad52 protein species form ring-structures, but lacks an NLS (blue).

Large nuclear proteins require active transport to enter the nucleus (Gorlich and Mattaj, 1996). If Rad52 forms a multimer in the cytosol, the Rad52-complex reaches a size preventing diffusion, but instead it requires active transport. It has been suggested that a protein containing several NLSs is more readily transported to the nucleus, and one could expect the same efficient transport of a heptameric ring-structure of Rad52 molecules

each expressing a weak NLS. The results presented therefore suggest a biological relevance of the ring structure formation of Rad52 (figure 26).

3. Identification of a novel Rad52 domain responsible for repair center assembly

One of the responses following induction of DNA DSBs is the formation of repair centers at break sites (Essers *et al.*, 2002), (Lisby *et al.*, 2003a). These repair centers are complexes containing several proteins including Rad52. Although the nature of the repair centers has been studied intensively in recent years, it is still unknown how and why Rad52 concentrates into these repair centers. In this study, as a means to determine regions in Rad52 important for this mechanism, a series of fluorescently tagged Rad52 mutants were constructed and their ability to form repair centers *in vivo* was investigated by using fluorescent microscopy. In this way, it was possible to isolate Rad52 mutants including internal alanine substitution mutants, which were unable to form repair centers. All mutant strains of this type showed severe defects in DNA DSBR as they were highly sensitive to MMS treatment. In an attempt to find an explanation for the phenotype of the mutant strain, a mutant protein was purified and several biochemical *in vitro* assays were conducted to test whether the mutation had any effect on other Rad52 functions in HR. The mutant protein was found to perform similar to wild-type Rad52 in the assays tested, thus indicating a more complex explanation for the mutant phenotype.

3.1 Introduction

In eukaryotes, DNA DSB repair is essential to maintain genetic stability, and the major pathway of DSB repair in *S. cerevisiae* is HR. Proteins involved in HR are encoded by members of the *RAD52* epistasis group, which includes *RAD50*, *RAD51*, *RAD52*, *RAD54*, *RAD55*, *RAD57*, *RAD59*, *RDH54*/*TID1*, *RFA1*, *MRE11*, and *XRS2*. Physical interactions between these proteins are central in HR processes. For instance, Mre11, Rad50 and Xrs2 form a complex that is involved in early events in DNA DSB repair including resection of the ends. Physical interactions have also been described for most of the other proteins in the group. Accordingly, Rad52 interacts with itself, Rad59, Rad51 and RPA. Rad51 interacts with itself, Rad52, Rad54 and Rad55, and Rad55 interacts with Rad57. Thus, the possibility of RPA, Rad51, Rad52, Rad54, Rad55, Rad57 and Rad59 to form a large protein network exists, which is depicted in figure 27 (Hays *et al.*, 1995), (Davis and

Symington, 2003). However, individual protein-protein interactions may be established and disrupted at different stages of the repair process.

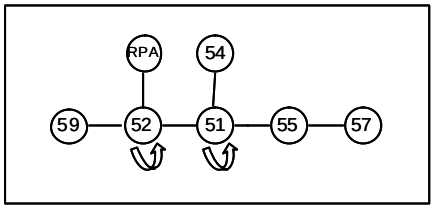


Figure 27. Schematic representation of possible protein-protein interactions in a repair center. Rad52 interacts physically with itself (indicated by the arrow), Rad59, RPA and Rad51. Rad51 furthermore interacts with itself as well as Rad54 and Rad55. Rad55 also binds physically to Rad57.

The biochemical importance of the physical interactions has been clarified in many cases, for example it has been firmly established that RPA, Rad51 and Rad52 collaborate to ensure an efficient strand invasion reaction (Sung, 1994), (New *et al.*, 1998), (Sugiyama and Kowalczykowski, 2002). Specifically, when the ends at a DNA DSB have been resected, the resulting single-stranded tails are covered by RPA to prevent formation of secondary DNA structures (Brill and Stillman, 1989), (Alani *et al.*, 1992). Next step in the repair process involves invasion of an intact homologous duplex catalyzed by the strand invasion activity of Rad51. However, *in vitro* it has been shown that single stranded DNA covered by RPA is inaccessible to Rad51 and recruitment of Rad51 requires the mediator function of Rad52 that interacts with both RPA and Rad51(Sung, 1997a), (New *et al.*, 1998). In living cells, the involvement of several repair proteins during repair can be visualized by determining the cellular re-localization of fluorescently tagged repair factors after introduction of DNA damage. Many of these proteins have been observed to concentrate at the site of DNA damage to form visible foci that may contain more than a thousand molecules of a single repair protein (Haaf *et al.*, 1995;Lisby *et al.*, 2001), (Miyazaki *et al.*, 2004). These foci, or repair centers, are highly dynamic structures, and the exchange rate of some proteins in a center is fast and for others it is slow (Essers *et al.*, 2002). In addition, some repair proteins are detected only at the early stages of the lifetime of a repair center, whereas other proteins may only enter at later stages of the repair process.

One protein that is known to be present in a repair center is Rad52 (Gasior *et al.*, 1998), (Liu *et al.*, 1999), (Lisby *et al.*, 2001;Essers *et al.*, 2002). In *S. cerevisiae* Rad52-YFP forms foci specifically during S-phase (Lisby *et al.*, 2001) indicating that repair is coupled

to replication. Interestingly, it has been determined that multiple DNA DSBs are located in a single repair center indicating that they are repaired simultaneously. Perhaps the ability of Rad52 to handle multiple lesions at the same time explains the large size of a repair center. Although much is known about the dynamics of repair proteins, it is not known what triggers a specific repair protein like Rad52 to enter a repair center. To understand the biological consequences of these repair centers, a mutagenesis study of *RAD52* was performed to isolate mutants, which failed to be incorporated into repair foci. In the present study several *rad52* mutations have been introduced, which causes a defect in Rad52 focus formation and the mutant fails to form foci spontaneously as well as when DNA DSBs are introduced by DNA damaging agents. One mutation was selected for further characterization and the mutated protein was purified and analyzed biochemically.

3.2 Materials and methods

3.2.1 Yeast strains and media

[a]Strain	Genotype
W1588-4C	MATa ade2-1 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 RAD52
UM74-3B	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 RAD52-YFP
J883	MATa ade2-1 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52-?327
UM94-9C	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52-?327-YFP
UM94-2D	MATa ADE2 his3-11,15 leu2-3,112 1 ura3-1 trp1-1 LYS2 rad52-?307-YFP
UM90-2C	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp 1-1 LYS2 rad52-?287-YFP
UM93-12D	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp 1-1 LYS2 rad52-?267-YFP
UM177-7E	MATa ade2-1 his3-11,15 leu2-3,112 ura3-1 trp1 LYS2 rad52-Q308A-D309A-D310A-D311A
UMR101-15B	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52::HIS5
UM68-5B	MATa ade2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52::HIS5

[a]All strains are derivatives of W303-1A and W303-1B (Thomas, B. J. & Rothstein, R. 1989).
In addition to the genotype listed above all strains are RAD5

Table 3. Yeast strains used in this study.

3.2.2 *RAD52* constructs

- *YFP* tagged *rad52* alleles

RAD52 mutants were constructed and fused to YFP using the cloning-free PCR-based allele replacement method previously described (Erdeniz *et al.*, 1997;Lisby *et al.*, 2001). PCR was performed using the Expand^{TM} high fidelity system according to the supplier's instructions (Roche Diagnostics). Rad52 truncations and YFP-fusions were made using

plasmids pRS413-*RAD52*, pWJ1164 and pWJ1165 as templates (Lisby *et al.*, 2001). Primers for *rad52-D267-YFP*, *rad52-D287-YFP* and *rad52-D307-YFP* encoding 267, 287 and 307 amino acid residues respectively were (Rad52start-F) 5' - ATGGCGTTTTTAAGCTATT, (Rad52_267_reverse_A) 5' - agttcttctcctttactcatTGCGGCCATCATTGGGGTAC, (Rad52_287_reverse_A) 5' - agttcttctcctttactcatATCGAGAGATTTGAGATCAGTATC, and (Rad52_307_reverse_A) 5'- agttcttctcctttactcat-AAAATCATCGCTAAACATAAGAG. Shown in small letters are adaptamer sequences complementary to the first 20 nucleotides of *YFP*.

The downstream *RAD52* fragment was amplified using primers (Rad52Tdown) 5' - GGATGAACTATACAAATAACCCGCTTCCTGGCCGAAAC and (Rad52_down) 5' - AATGAACCTAAGGATTCCGC. Primer sets (GFPstart-F) + (3'int) 5' - GAGCAATGAACCCAATAACGAAATC and (GFPTdown) 5' - TCCCCGCGGTTATTTGTATAGTTCATCC + (5'int) 5' - CTTGACGTTCGTTCGACTGATGAGC were used to amplify the *YFP-URA3* fragments.

- Alanine substitutions in *RAD52*

Plasmid, pWJ1213 (kind gift from Michael Lisby) was used as template to construct the *RAD52* mutant vectors. pWJ1213 harbors *RAD52* fused C-terminally to YFP. A *Xma*I and *Sac*I fragment was removed from pWJ1213 in order to create unique enzyme restriction sites flanking the *RAD52* region of interest. T4 polymerase (New England Biolabs) was used to remove 3' overhangs at the *Sac*I site and to fill in the 5' overhangs at *Xma*I cut site before religation. T4 DNA ligase (New England Biolabs) was used for religating the plasmid, according to supplier's instructions.

The resulting plasmid, pWJ1213-*?XmaI-SacI* was used to clone the inserts that contained the alanine substitution sites for *RAD52*. The region between amino acid residues 299 to 311 was mutated by changing nine charged amino acid codons in the gene (from DSLMFSDDFQDDD to AALMFAAAFAAAA). In order to make up recognition sites for *XbaI* and *PacI* for cloning, the codons adjacent to the mutation region LLD (aa 296-298) and LIN (aa 312-314) were changed without changing the corresponding amino acid residue. CTT CTC GAT was changed to CTT CTA GAT and TTG ATA AAT was changed to TTA ATT AAT. The substitution of charged amino acid residues with alanines was done by PCR amplification. During the first round of PCR, (Nterm-seq4) 5' -

GTGGAGAACGAAAGACGGAAACC and (Ala_rv_primer) 5' - ATTAATTAAAGCGGCTGCCGCAAAGGCCGCTGCAAACATAAGGGCAGCATCTAGAAG GTCATCTTGATCCTGTTTGG were used to amplify the region with the *Xba*I recognition site to create fragment A. (Rad52-down) 5' - AATGAACCTAAGGATTCCGC and (Ala_fw_primer) 5' - CTTCTAGATGCTGCCCTTATGTTTGCAGCGGCCTTTGCGGCAGCCGCTTTAATTAATATGGGC AACACAAACAGTAA were used to amplify the site that contained a *PacI* site to create fragment B. Ala_fw_primer and Ala_rv_primer contained the sequences of the nine alanines that were designed for internal substitution. In a second round of PCR, fragments A and B were fused to become insert C. Insert C was cloned into pWJ1213-*?XmaI-SacI* with *BbvC*I and *Sph*I.

Primers used for the construction of plasmids pRad52-2Ala, pRad52-3Ala and pRad52-4Ala were (2ala_fw) 5' - CTAGATGCTGCCCTTATGTTTAGCGATGATTTTCAAGACGACGACTTAAT, (2ala_rv) 5' - TAAGTCGTCGTCTTGAAAATCATCGCTAAACATAAGGGCAGCAT, (3ala_fw) 5' - CTAGATGATTCTCTTATGTTTGCAGCGGCCTTTCAAGACGACGACTTAAT, (3ala_rv) 5' - TAAGTCGTCGTCTTGAAAGGCCGCTGCAAACATAAGAGAATCAT, (4ala_fw) 5' - CTAGATGATTCTCTTATGTTTAGCGATGATTTTGCGGCAGCCGCTTTAAT, and (4ala_rv) 5' – TAAAGCGGCTGCCGCAAAATCATCGCTAAACATAAGAGAATCAT, respectively. The designed primers, together with the primers of complementary sequences were annealed at 98ºC for 2 minutes with 10x annealing buffers (10X Annealing Buffer: 200 mM Tris-HCl, pH 7.5, 500 mM NaCl, 100 mM $MgCl_2$). The tubes with the primer mixtures were allowed to cool until room temperature was reached. The annealed fragments were then used for ligation with the mutant vector and transformed into *E. coli*. Purification of DNA during vector construction was done using the GFX^{TM} PCR DNA and Gel Band Purification Kit (Amersham Pharmacia Biotech).

3.2.3 Microscopy

In experiments with integrated *rad52* mutations cells were grown in SC medium and in the experiment with *rad52* mutants expressed from plasmid cells were grown in SC-His medium to select for the plasmid. In all experiments, cells were grown at 23°C (to allow the

YFP chromophore to form efficiently) in 3 ml overnight cultures to an OD_{600} of 0.3. The overnight cultures were spun down and resuspended in 0.5% MMS (M4016 from Sigma), and then incubated for 15 min. Next, cultures were washed twice with water and resuspended in fresh SC or SC-His medium as appropriate. The cells were then incubated for 30 min. to allow Rad52-YFP foci to form before imaging. For all epifluorescence microscopy, cells were spun down and immobilized on a glass slide by mixing them with a 37°C solution of 1.2% (wt/vol) low melting agarose (NuSieve 3:1 from FMC) containing the appropriate medium.

Live cell images were captured with a cooled Evolution QEi monochrome digital camera (Media Cybernetics Inc., USA) mounted on a Nikon Eclipse E1000 camera (Nikon, Japan). Images were captured at 100-fold magnification using a Plan-Fluor 100x 1.30 NA objective lens. The illumination source was a 103W mercury arc lamp (Osram, Germany). The fluorophore YFP was visualized using a band-pass YFP filter (EX500/20, DM515, BA520 combination filter, Nikon, Japan). Exposure time for Rad52-YFP was 1.5 sec. with a 12.5% neutral density (ND8) filter in place to reduce photobleaching. For each field of cells, nine to 11 fluorescent images were obtained at 0.4 µm intervals along the Z-axis to allow inspection of all focal planes of each cell.

Small budded cells were recognized as cells where the daughter cell was less than 30% of the mother, in large budded cells, the daughter cell was larger than 30% of the mother and no budded cells had no bud or, if the cells were together, their nucleus were totally separated from each other.

3.2.4 MMS spot assay

The mutagenized plasmids were transformed into *rad52D* (UMR101-15B), and cells were grown to stationary phase in SC-His media at 30°C and resuspended in 50 mM potassium phosphate buffer, pH 7.0. At approx. 1 x 10^8 cells/ml, cell cultures were incubated for 0 or 20 minutes with 0.5% MMS, and the reaction was stopped with 1 volume of cold 10% sodium thiosulfate – $5H_20$. Appropriate dilutions were made and viability was determined after spotting on solid SC-His medium. Plates were incubated at 30°C for 2 days before examination (Prakash and Prakash, 1977a).

3.2.5 *RAD51* overexpression

The series of *rad52* mutant strains were transformed with the galactose inducible *RAD51* overexpression plasmid, pYESS10Rad51 (Jiang *et al.*, 1996) as well as an empty vector, for control. Cell cultures were grown in liquid SC-Ura medium containing 2% galactose to a cell density of 10^7 cells/ml. Cells were then resuspended in 50 mM KH_2PO_4 buffer (pH7) and treated with 0.5% MMS and incubated with shaking at 30^o C. 0.75 ml of each sample was withdrawn at 0, 10 and 15 minutes (cells that were withdrawn at 0 minute did not contain MMS and served as control). The reaction was stopped with 10% cold $Na_2S_2O_3$ (Prakash and Prakash, 1977b) and samples were kept on ice at all time. The OD_{600} of all withdrawn samples were adjusted to 0.3 (which corresponds to cell density of 10^7 cells/ml), except for the *rad52* null strain where the OD was adjusted to 0.6. Spot assays were carried out as previously described in this report, except that the cell cultures were spotted onto uracil deficient agar plates containing 2% galactose.

3.2.6 Purification of Rad52 and Rad52- 4Ala

Amino acid residues 308-311 in wild-type Rad52 were substituted with alanine in the expression vector pET-11d-RAD52-6his (a kind gift from Patrick Sung, to be described at a later time) by site-directed mutagenesis according to the supplier's instructions (Stratagene). The plasmid encodes *S. cerevisiae* Rad52 protein, starting from the 34^{th} N-terminal amino acid residue of the protein with six histidine tags fused to the C-terminus. The oligo sequences were the following: 5' CTTATGTTTAGCGATGATTTTgctgccgcagcgTTGATAAATATGGGCAACAC and 5' GTGTTGCCCATATTTATCAAcgctgcggcagcAAAATCATCGCTAAACATAAG (the codon substitution is shown in small letters). pET11d-Rad52-Q308A-D309A-D310A-D311A-His_6 together with pET-11d-RAD52-6his were introduced into *E. coli* strain Rosetta (Novagen), and expression was induced by 0.4 mM isopropyl ß-D-thiogalactoside (IPTG) and grown at 37°C for 3 hours. Purification steps were carried out at 4°C. Lysate was prepared from *E. coli* cell paste using a French press in breaking buffer (100 mM Tris-HCl, 20% sucrose, 4 mM EDTA, pH 7.5 containing 200 mM KCl and the protease inhibitors aprotinin, chymostatin, leupeptin, and pepstatin A at 5 mg/ml each, as well as 1 mM Benzamidine, 1

mM ßMer, 0.01% IgePal and 0.5 mM PmSF). The crude lysate was clarified by centrifugation (40K, 60 min, 4C), and the supernatant from the centrifugation step was applied to glutathione-Sepharose 4B (Amersham Biosciences) followed by Ni-NTA beads (Qiagen). The beads were washed three times with buffer K + 500mM KCl (buffer K: 20 mM KH_2PO_4, pH 7.4, 150 mM KCl, 10% glycerol, 0.01% Nonidet P-40) containing 10 mM, 20 mM and 30 mM Imidazole respectively. Bound His-fusion protein was eluted with 200 mM Imidazole in buffer K+ 500 mM KCl. The final protein pools were concentrated to 50 µl in a Centricon-30 device (Amicon) and stored in 2 µl aliquots at – 80° C.

3.2.7 DNA binding assays

The DNA binding experiments were carried out with ssDNA 80mer ATGAACATAATTGAAATAAGGATCCGGCTAATACAAAATAAGTAAAAGGTTAAACATAG AATTCAAAGTAAAGGATATAA and dsDNA 40mer (forward TTATATCCTTTACTTTGAATTCTATGTTTAACCTTTTACT. The FX 174 viral (+) strand was purchased from New England Biolabs, and the replicative form (about 90% supercoiled form and 10% nicked circular form) was from Invitrogen. These oligonucleotides were purified from a 15% polyacrylamide gel as described previously (Trujilo and Sung). The two oligonucleotides were labeled with [?^{32}P] ATP and T4 polynucleotide kinase (New England Biolabs). Varying amounts of Rad52 or Rad52-4Ala (29 – 450 nM) protein was incubated with ^{32}P-labeled 40mer (10 nM nucleotides) and 80mer (10 nM nucleotides) at 37°C in 10 µl of buffer (40 mM Tris-HCl, pH 7.8, 50 mM KCl, 1 mM DTT, and 100 µg/ml bovine serum albumin) for 15 min. To release the bound DNA, the reaction mixture was deproteinized with 0.5% SDS and 500 µg/ml proteinase K at 37°C for 5 min. After addition of gel loading buffer (50% glycerol, 20 mM Tris-HCl, pH 7.4, 2 mM EDTA, 0.05% orange G), the reaction mixtures were resolved in 10% native polyacrylamide gels at 4°C in TAE buffer (40 mM Tris-HCl, pH 7.4, 0.5 mM EDTA) and dried before subjected to phosphorimaging analysis.

3.2.8 Single stranded DNA annealing assays

Single stranded annealing of a ^{32}P labeled 40-mer (oligo 1) 5´-ATTAAGCTCTAAGCCATGAATTCAAATGACCTCTTATCAA and a complementary

unlabeled 80-mer oligo with 40 base pair overhang (oligo 2); 5' - TTGATAAGAGGTCATTTGAATTCATGGCTTAGAGCTTAATTGCTGAATCTGG TGCTGGGATCCAACATGTTTTAAATATG. Oligo 1 (6 nM nucleotides) and radio labeled Oligo 2 (6 nM nucleotides) were mixed and added to Rad52 or Rad52-4Ala (13 nM – 85 nM) to start the reaction. The completed reactions (15 µl) were incubated at 25 °C, for 3 minutes, 9 µl of the annealing reactions was removed and treated with 0.5% SDS, 500 µg/ml proteinase K, and an excess of unlabeled Oligo-1 (40 nM nucleotides) at 30 °C for 5 min. The samples were resolved in 10% native polyacrylamide gels run in TAE buffer and dried before subjected to phosphorimaging analysis.

3.2.9 Pull down assays

Purified His-tagged Rad52 (4.17 µM) was incubated with purified Rad51 (a kind gift from Patrick Sung) (3.87 µM) in 30 µl buffer K + 300 mM KCl with 1 mM DTT and 0.01% IgePal and incubated at 4°C for 30 min before the reaction mixture was mixed with 7.5 µl of Nickel-NTA agarose beads (Qiagen) and incubated at 4°C for an additional 30 minutes. The beads were then washed twice with 400 µl of the same buffer with 150 mM KCl, and the bound proteins were eluted with 30 µl of 3% SDS. The supernatant that contained unbound protein, the KCl wash, and the eluate were run on a 12% SDS-PAGE gel and subsequently stained with Coomassie Blue.

3.2.10 Gel filtration assay

A 15 ml Sepharyl^{TM} S-300 High Resolution bead gel filtration column (Amersham Pharmacia Biotech AB) was equilibrated overnight in buffer K+150 mM KCl with 0.01% IgePal and 1 mM DTT at 0.1 ml/min. Gel filtration standard mix from BioRad (diluted with 50 µl K+150) was loaded on the column, and the column was washed 10x with 100 ul K+150. Next, purified Rad52 or Rad52-4Ala (80 µg) was loaded on the column and fractions collected by a FPLC (0.1 ml/min with buffer K + 150 KCl). Samples were run on 12% SDS gel and stained with Coomassie Blue.

3.2.11 Strand exchange assay

The standard DNA strand exchange reaction was performed as described by Sung *et al*. (Sung, 1997a). Briefly, Rad51 (0 - 10 µM) was incubated with FX174 ssDNA (30 µM nucleotides) in 10 µl of buffer (150 mM K-MOPS, pH 7.2, 5 mM DTT, 50 mM KCl, 2.5 mM ATP, and 3 mM $MgCl_2$) for 5 min. After addition of RPA, the reaction mixtures were incubated for another 5 min before the incorporation of 1 µl of double-stranded DNA (25 µM nucleotides) and 1 µl of 50 mM spermidine hydrochloride. At times 50 and 80 min. a 4.5 µl portion of the reaction mixtures was withdrawn, mixed with 0.1% SDS and kept on ice. Samples were deproteinized with proteinase K for 15 min. at 37° C, and then analyzed in 0.9% agarose gels in TAE buffer. The gels were treated with EtBr to stain the DNA species. The RPA inhibition assay was performed like the above, except that the ssDNA was incubated both with Rad51 and RPA for 5 minutes before dsDNA was added. The concentration of RPA used was 4.17 µM – 6.96 µM. The strand exchange reaction restoration reaction was performed similarly, except that the ssDNA was incubated with Rad51 and RPA and increasing concentrations of Rad52 (0.57µM – 1.13 µM).

3.3 Results

Rad52 is a multi-domain protein and several functional regions in the protein have been mapped, including those that are involved in Rad52 self-association, DNA binding and in forming interactions with Rad51, RPA and Rad59 (Figure 28).

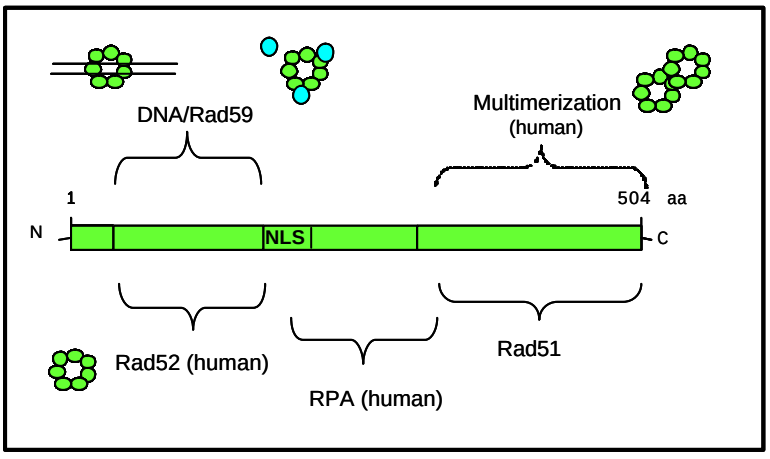


Figure 28. Functional map of Rad52. Rad52 has multiple functions including binding DNA and several proteins including Rad59, RPA, Rad51 and itself. A putative NLS introduced in the previous chapter is also shown on the map. Functional domains identified in human Rad52 are indicated in brackets.

In vivo, Rad52 is recruited to a repair center following DNA damage, but no region in Rad52 has been assigned for this function. To investigate the possibility that a specific region in Rad52 is required for its recruitment to a repair center, a mutational strategy was employed. Initially, a series of Rad52 deletion mutations were constructed where the protein was progressively shortened from the C-terminus (Figure 29).

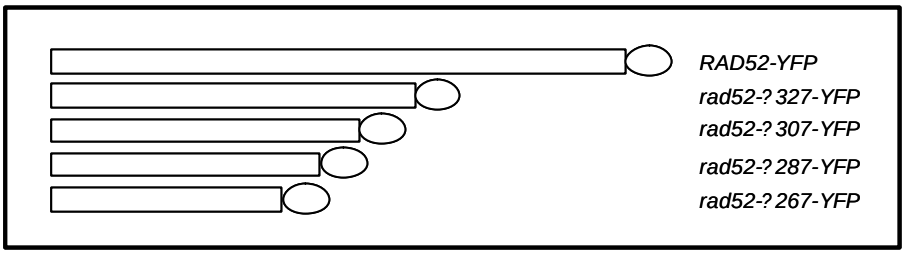


Figure 29. Truncated *RAD52* alleles used in this study. The relative size of the constructs is indicated. The *RAD52* constructs are fused to *YFP* at the C-terminal.

The C-termini of all Rad52 truncations were extended by YFP to allow detection of the cellular distribution of the resulting fusion proteins by fluorescence microscopy. Accordingly, the ability of the mutant strains to develop a mutant-Rad52-YFP derived focus, during repair of spontaneous as well as during MMS-induced damage could be determined. The Rad52-YFP fusions were confirmed by Western blotting (data not shown).

3.3.1 KlRad52-YFP can form foci in *S. cerevisiae*

In the previous section it was reported how *K. lactis* Rad52 can sort to the nucleus in a *S. cerevisiae rad52D* strain, so the first experiment to conduct was to investigate whether the Rad52 focus formation property is conserved among the two yeast species as well. KlRad52-YFP (construction described in the previous chapter) was transformed into a *S. cerevisiae rad52D* strain and transformants were treated with MMS and subsequently examined by fluorescent microscopy. Like ScRad52-YFP, KlRad52-YFP formed foci after DNA damage induction (figure 30).

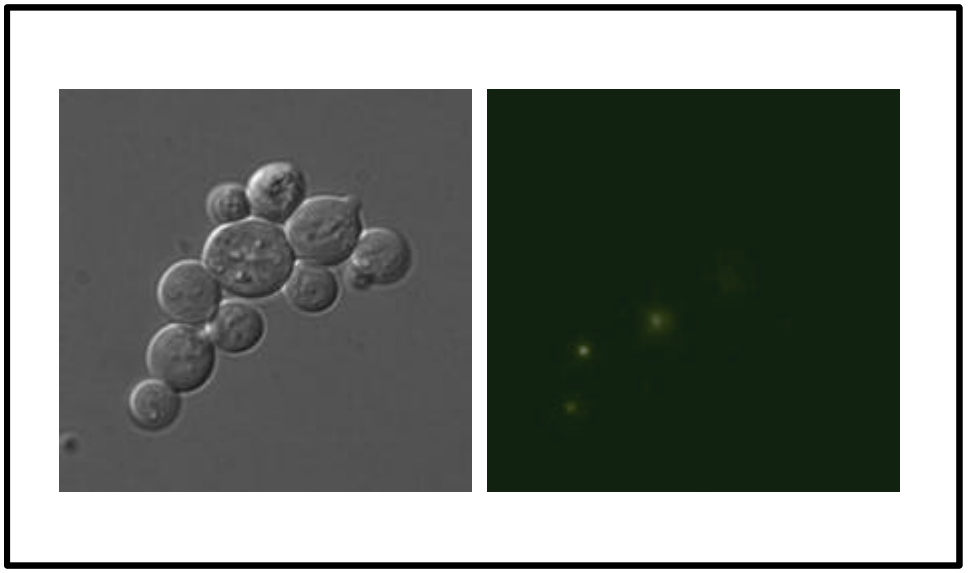


Figure 30. KlRad52-YFP expressed in *S. cerevisiae*. CEN based vector harboring KlRad52-YFP is transformed into *S. cerevisiae rad52D* strain and cells treated with 0.5 % MMS for 15 minutes before subjected to fluorescent microscopy.

Still, it is unknown if the Rad52-YFP foci represent active repair, but the result is in agreement with observations by Milne and Weaver who found KlRad52 to partly complement the DNA damage sensitivity of a *S. cerevisiae rad52D* strain (Milne and Weaver, 1993). There are many stretches of homology between the sequences of ScRad52 and KlRad52, so to narrow down the important region for repair center formation of ScRad52, a mutational strategy was again applied and a series of Rad52 truncations were constructed and investigated for their ability to form repair centers with and without MMS treatment.

3.3.2 Rad52 mutant species form more repair centers

First, the ability of wild type Rad52-YFP to form repair centers was compared to the results obtained previously (Lisby *et al.*, 2001). The number of wild-type cells containing spontaneous foci was lower compared to what previously has been reported (Lisby *et al.*, 2001). The difference may be due to small changes, e.g. growth conditions, between how the assay is conducted between laboratories. However, when cells were treated with MMS a large number of cells containing repair centers was observed and the number was similar to the number observed by Lisby *et al*. using γ-ray (Lisby *et al.*, 2001). This comparison was done to ensure the functionality of the Rad52-YFP assay. Accordingly, the Rad52-YFP assay presented here is clearly functional.

Next, the ability of the *rad52-D327-YFP* strain to form foci was investigated. In the collection of Rad52 truncation mutants it encodes the species with the smallest Rad52 deletion. With this strain, the number of cells containing foci before and after MMS treatment was dramatically increased compared to cells expressing wild-type Rad52-YFP. Specifically, in strains expressing Rad52-Δ327-YFP, the numbers of cells containing at least one spontaneous focus in small and large budded cells were increased 30-fold and 8-fold, respectively (table 4).

Genotype	Spontaneous foci (% of cells)			MMS induced foci (% of cells)		
	Un-budded cells	Small budded cells	Large budded cells	Un-budded cells	Small budded cells	Large budded cells
RAD52-YFP	0	1	3	44	81	69
rad52-?327-YFP	2	31	24	81	94	88
rad52?307-YFP	0	4	2	0	0	5
rad52?287-YFP	0	4	0	0	0	3
rad52?267-YFP	0	0	0	0	4	2
rad52?237-YFP	0	0	0	0	0	0

Table 4. Focus formation of *rad52* truncated strains. *rad52D*cells are transformed with the mutagenized plasmids and inspected by fluorescent microscopy before and after MMS treatment and the number of cells with foci is established. A minimum of 100 cells of each strain is investigated.

Moreover, a spontaneous focus was occasionally observed in un-budded *rad52-D327-YFP* cells, but never in un-budded *RAD52-YFP* cells (Table 4). After induction of DNA damage by MMS, most *rad52-D327-YFP* cells contained foci. In fact, for all cell types, the number of cells with a Rad52 focus was higher with *rad52-D327-YFP* strain compared to the numbers observed with wild-type *RAD52-YFP* cells. The increase of cells containing spontaneous as well as MMS-induced foci in *rad52-D327-YFP* strains compared to wild-type strains could be caused by an accumulation of unrepaired lesions in *Rad52-D327* cells due to the inability of Rad52-Δ327-YFP to interact with Rad51. The largest difference, two-fold, was observed with un-budded cells.

3.3.3 Rad52 mutants unable to form repair centers

The remaining Rad52 truncation species, which span truncations that terminate at aa residues 237 to 307, were examined for their localization in the nucleus in the presence and absence of MMS. Interestingly, the number of cells containing spontaneous repair centers is severely reduced in these mutant strains compared to *rad52-D327-YFP* strains.

In fact, hardly any cells containing a focus were observed (table 4). A similar picture was observed when MMS treated cells were inspected. Hence, the number of cells containing at least one MMS induced focus was more than ten-fold reduced in strains expressing one of these Rad52 truncations compared to the numbers obtained for all cell types with *rad52-D327-YFP* as well as with wild-type *RAD52-YFP* strains.

3.3.4 MMS sensitivity of Rad52 mutants suppressed by *RAD51* overexpression

Strains that express truncated Rad52 species that terminate in the amino acid residue range of 169 to 327, have been shown to be MMS sensitive (Asleson *et al.*, 1999). Some of these strains, like *rad52-D292* and *rad52-D327*, can be fully or partially suppressed by overexpression of *RAD51*. *rad52-D169* is not suppressed, but *rad52-D210* exhibit a slightly higher survival when *RAD51* is overexpressed. Considering the size of YFP, the possibility existed that the strains, *rad52-D267-YFP*, *rad52-D287-YFP* and *rad52-D307-YFP*, fail to form foci due to the influence of YFP. To investigate this possibility, it was tested whether overexpression of *RAD51* could suppress the MMS sensitivity of the strains expressing the Rad52 truncation species ranging from Rad52-Δ267-YFP to Rad52-Δ307-YFP. *RAD52* strains were individually transformed with a high copy plasmid expressing *RAD51* and with an empty vector for control and spotted on solid plates with and without MMS. Next, the survival of the strains was inspected and in all cases, overexpression of *RAD51* suppressed the MMS sensitivity of the mutant strains 10 – 100 fold, (figure 31) to levels that are comparable to those obtained with strains expressing similar, but untagged, Rad52. It was therefore concluded that the failure of these Rad52 truncations to form foci are likely due to a missing Rad52 function rather than to obstruction caused by the presence of YFP.

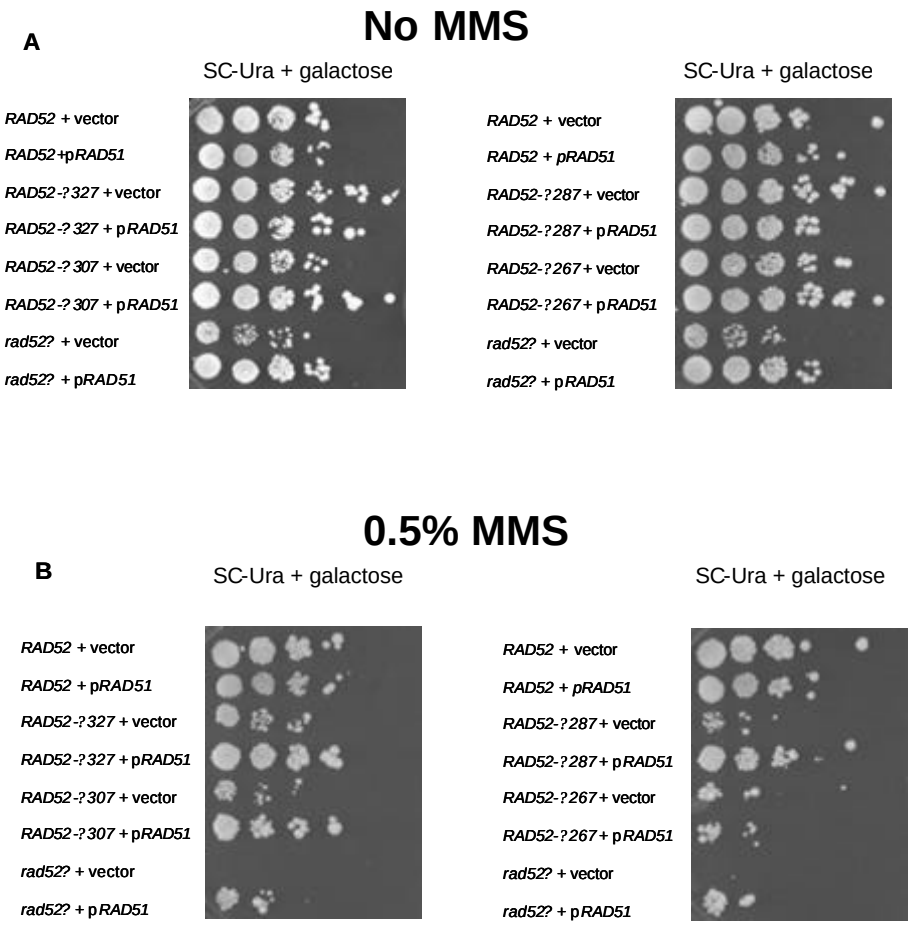


Figure 31. Overexpression of *RAD51*. (**A**) *RAD52-YFP, rad52-D327-YFP, rad52-D307-YFP, rad52-D287-YFP, rad52-D267-YFP and rad52D* strains transformed with *RAD51* overexpression plasmid (pRAD51) or empty vector (vector) spottet on SC-Ura plates with galactose. (**B**) *RAD52-YFP, rad52-D327-YFP, rad52-D307-YFP, rad52-D287-YFP, rad52-D267-YFP and rad52D* strains transformed with *RAD51* overexpression plasmid or empty vector and treated with 0.5 % MMS for 10 min and spotted on SC-Ura plates containing galactose.

3.3.5 The middle part of Rad52 is important for repair center formation

The observation that strains expressing Rad52-Δ327-YFP, but not those that express Rad52-Δ307-YFP, form foci suggest that a region in Rad52 important for Rad52 repair center formation is present between or in the vicinity of aa residues 307 to 327. As discussed earlier, Rad52-YFP from *K. lactis* re-localizes and forms foci in *S. cerevisiae rad52D* cells and the relevant amino acid sequence section of *S. cerevisiae* Rad52 was

therefore compared to the corresponding section of *K. lactis* Rad52 (figure 32). Interestingly, a number of aa residues are present in both Rad52 species in a region that spans aa residues 296 - 314. Inspired by this finding, it was decided to construct a plasmid expressing a *S. cerevisiae* Rad52-YFP species where nine of these highly charged aa residues simultaneously were replaced with alanine.

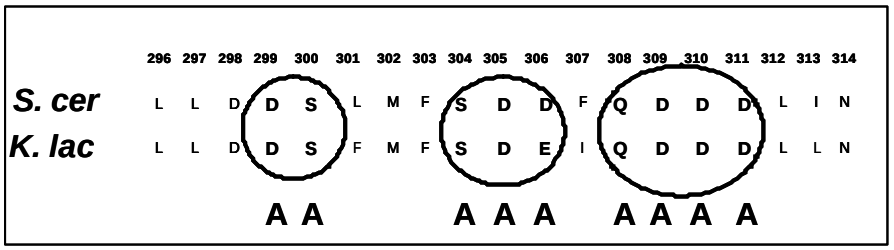


Figure 32. Nine charged amino acid residues are substituted with alanine in conserved region of ScRad52.

This plasmid was transformed into a *rad52D* strain and the ability of the Rad52 mutant Rad52-D299A-S300A-S304S-D305A-D306A-Q308A-D309A-D310A-D311A-YFP protein (hereafter called Rad52-9Ala-YFP) to form repair foci in the presence and absence of MMS was determined and compared to the corresponding numbers obtained with strains expressing wild-type *RAD52-YFP* from a plasmid (the two top rows in table 5).

genotype	Spontaneous foci (% of cells)			MMS induced foci (% of cells)		
	Un-budded cells	Small budded cells	Large budded cells	Un-budded cells	Small budded cells	Large budded cells
RAD52-YFP	6	19	36	25	61	76
rad52-9Ala-YFP	2	11	6	4	4	26
rad52-4Ala-YFP	1	0	2	1	0	4
rad52-3Ala-YFP	0	0	0	2	0	1
rad52-2Ala-YFP	2	14	7	4	11	15

Table 5. Focus formation of YFP-tagged Rad52 proteins. *rad52D* cells are transformed with the mutagenized plasmids and inspected by fluorescent microscopy before and after MMS treatment and the number of cells with foci is established. A minimum of 500 cells of each strain is investigated.

It was noticed that the number of cells containing a spontaneous Rad52 repair center was higher in strains expressing wild-type Rad52-YFP from a plasmid compared to strains expressing it from the endogenous *RAD52* locus (compare top left panel of table 4 and 5). Perhaps the difference is due to the slower metabolic rate, which results from growth in selective medium compared to growth in rich medium. Hence, if the turnover of Rad52 repair centers is slowed down, cells containing Rad52 foci will accumulate and constitute a larger fraction of the total population of cells. This effect should be reduced when damage

is induced in all cells in a population simultaneously. In agreement with this, the number of cells containing repair centers after MMS treatment was more similar when strains grown in selective- and rich media were compared. Compare top right panel in table 4 with the top right panel in table 5.

A high number of cells was observed to contain spontaneous wild-type Rad52-YFP foci in selective medium. The Rad52-9Ala-YFP mutant did not form as many spontaneous foci. The number of foci is 3-fold lower in un-budded cells and 6-fold lower in large budded cells. The results obtained for Rad52-9Ala-YFP are comparable to the results obtained with the strain deleted in this region, *rad52D307-YFP*. It strongly indicates that the failure of Rad52-Δ307-YFP to be recruited to sites of DNA damage is due to loss of a Rad52 function present in the area between or in the vicinity of aa residues 307-327 rather than to gross mis-folding of this protein species.

To further investigate which of the nine amino acid residues identified above are required for recruitment of Rad52 to DNA damage, plasmids expressing three additional mutant Rad52-YFP species were constructed. Specifically, the region encompassing the nine mutated amino acid residues was split into three new discrete regions and each area was mutated to contain a double, a triple and a quadruple amino acid residue substitution, respectively, indicated in rings on figure 32. Combined, the three new mutations covered all of the original nine residues analyzed above as shown in figure 32. After transformation, the cellular localization of the three fluorescently tagged Rad52 mutants was determined. Of the resulting strains, those that expressed Rad52-D299A-S300A-YFP (Rad52-2Ala) contained a substantial number of cells with foci (table 5). Specifically, in the populations of small budded and large budded cells the numbers of cells with foci were 1.4- and 5-fold reduced compared to the corresponding numbers obtained with cells expressing Rad52-YFP. After MMS treatment, the number of cells containing foci did not increase dramatically indicating that this mutant species forms foci at a reduced rate or that they fail to be recruited to MMS induced damage. Hardly any of the cells from strains expressing the two other Rad52 mutants, Rad52-S304A-D305A-D306A-YFP (Rad52-3Ala) or Rad52-Q308-D309A-D310A-D311A-YFP (Rad52-4Ala) contained any Rad52* foci both with and

without MMS treatment (table 5). It was concluded that the region between aa 299-311 is important for the recruitment of Rad52 to a repair center.

3.3.6 Rad52 mutation results in MMS sensitivity

To determine if Rad52-YFP focus formation is a prerequisite for efficient DNA DSB repair, survival following MMS treatment of the *rad52* mutants was tested. *rad52D* cells were transformed with plasmids harboring wild-type *RAD52-YFP, rad52-2Ala-YFP, rad52-3Ala-YFP, rad52-4Ala-YFP, rad52-9Ala-YFP* or an empty vector. The transformants were treated with 0.5 % MMS for 0 or 10 min. and spottet on solid SC-His plates (figure 33). *RAD52-YFP* was only weakly sensitive to MMS whereas all the mutated strains showed low survival (figure 33). None of the mutations resulted in a dominant negative effect.

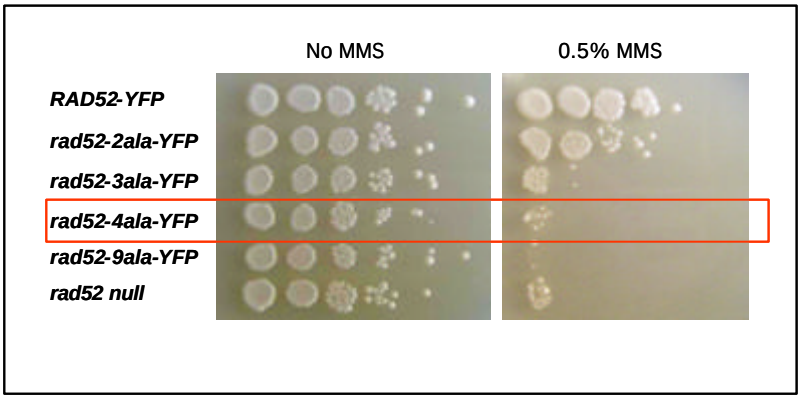


Figure 33. MMS spot assay with *rad52* alleles substituted with alanine residues in conserved region. The mutagenized plasmids are expressed in a *rad52D* strain, treated with 0.5 % in either 0 min or 10 min and spotted onto SC-His plates. The box indicates the *rad52-4Ala-YFP* mutant, which displays the most severe phenotype among the three mutants *rad52-2Ala*, *rad52-3Ala* and *rad52-4Ala*.

rad52-2Ala was mildly sensitive to MMS; the survival was reduced 10-fold when compared to wild-type *RAD52.* More severe sensitivity was seen for the *rad52-3Ala* and *rad52-4Ala* strains, which both displayed severe phenotypes with 1000 and 10000 fold reduced survival, respectively. Like *rad52-4Ala* the mutant allele with nine alanines substituted showed poor survival similar to a *rad52D* strain (figure 33). The mutant strain expressing the Rad52 species containing four alanine substitutions produced the most severe phenotype of the smaller substitutions and was therefore the chosen candidate for further characterization.

3.3.7 Purification of Rad52 and Rad52-4Ala

The biological assays presented above suggest that the amino acid residues 308 - 311 in Rad52 are required for repair center formation of Rad52 as well as for the DNA DSBR of MMS induced damage. The molecular basis for this phenotype could perhaps be clarified by characterizing the biochemistry of the Rad52-4Ala mutant in biochemical assays. Thus it was tested whether any of the known Rad52 functions are impaired in the mutant protein, Rad52-4Ala.

First, His-fusion proteins, Rad52 and Rad52-4Ala were expressed from pET-11d-based vectors in *E. coli* by IPTG induction and purified using affinity chromatography according to the purification scheme below (figure 34A). Both wild-type Rad52 and Rad52-4Ala proteins were purified to near homogeneity with no visible degradation products (figure 34B). In addition Western blot analysis verified that the purified proteins were indeed Rad52 (data not shown).

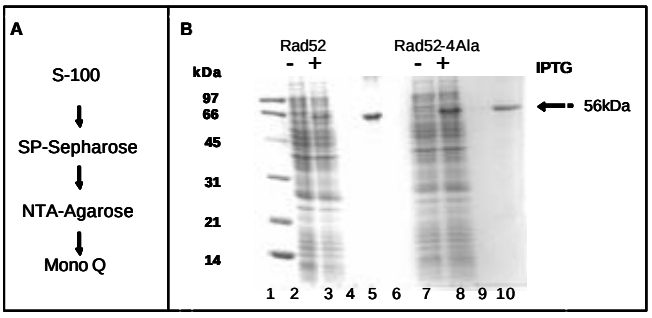


Figure 34. Purification of Rad52 and Rad52-4Ala. (A) The figure shows the purification scheme used for purifying Rad52 and Rad52-Ala. (B) Extracts from *E. coli* cells harboring pET-11d-RAD52-6his and pET-11d-rad52-4Ala-6his grown with (lane 3 and 8) and without (lane 2 and 7) IPTG and purified Rad52 (2µg) and Rad52-4Ala (2µg) (lane 5 and 10) are analyzed by denaturing polyacrylamide gel and and stained with Coomassie Blue.

3.3.8 Rad52 mutation does not affect DNA binding

As first reported by Mortensen *et al.* (Mortensen *et al.*, 1996) and later confirmed by others (Shinohara *et al.*, 1998), (Ranatunga *et al.*, 2001), (Kagawa *et al.*, 2001) Rad52 binds ssDNA and dsDNA. Thus, the DNA binding capacity of Rad52-4Ala was investigated and compared to that of wild type Rad52 in a competition assay. Increasing concentrations of Rad52 and Rad52-4Ala protein was incubated with radio labelled ss- and dsDNA substrates. Both proteins bound ss- as well as dsDNA. Figure 35 shows that the DNA binding properties of Rad52 is also present in Rad52-4Ala since the ss- and the dsDNA bands shift in the gel in the presence of both Rad52 and Rad52-4Ala.

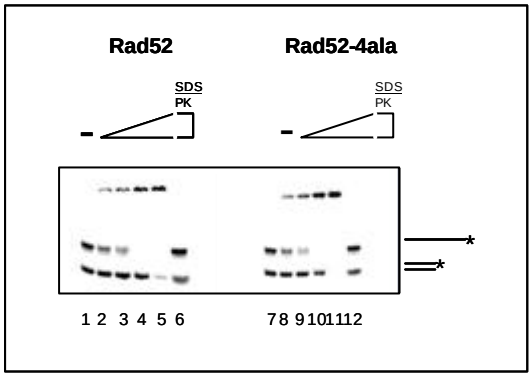


Figure 35. DNA binding assay. Increasing concentrations of Rad52 and Rad52-4Ala (29 – 450 nM) are incubated at 37° C for 15 min with a radio labeled 40-mer (10nM) and a radio labeled 80-mer (10nm). Reactions are terminated with 0.5 % SDS, 500 µg/ml proteinase K at 37° C for 5 minutes. Lanes 1 and 7 serve as negative controls with no protein added. Lanes 6 and 12 are controls to show that the proteinized proteins do not bind DNA. The reaction mixtures are resolved in a 10 % non-denaturing polyacrylamide gel that is dried and subjected to phosphorimaging analysis.

3.3.9 Rad52 mutation does not affect single strand annealing

Rad52 is uniquely required for the single strand annealing pathway of HR. Consistent with this genetic requirement, purified Rad52 can anneal complementary DNA strands (Mortensen *et al.*, 1996). The next question asked was therefore whether the Rad52 mutant Rad52-4Ala can anneal complementary ssDNA strands by using the assay described by Krejci *et al*. (Krejci *et al.*, 2002). The assay showed that both Rad52 and Rad52-4Ala exhibit DNA annealing activity and no difference in DNA annealing efficiency between Rad52 and Rad52-4Ala was detectable (figure 36). The annealing process was however inhibited at high concentration of Rad52 and Rad52-4Ala. The ssDNA annealed to dsDNA even without protein present, but was significantly more effective when protein was added to the reaction.

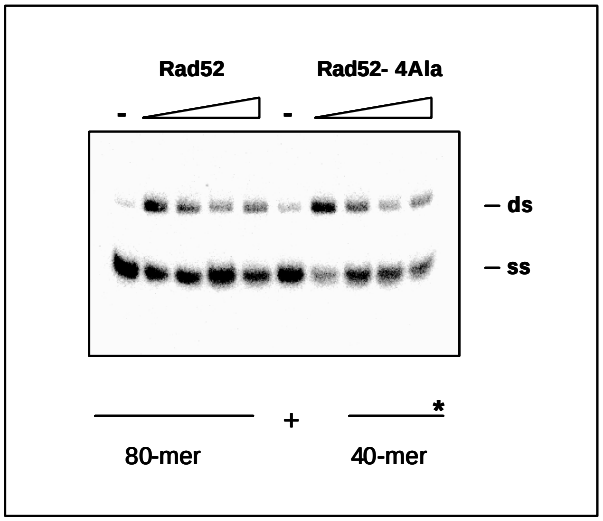


Figure 36. ssDNA annealing assay. A radio labeled 40-mer (6 nM) and a complementary unlabelled 80-mer oligo with 40 base overhang (6 nM) is incubated for 3 min. at 25° C with increasing concentrations of Rad52 and Rad52-4Ala (13 – 85 nM) and the reactions are terminated by treatment with unlabelled 40-mer, 0.5% SDS and 500 µg/ml proteinase K for 2 min. at 37° C. The reaction is also conducted without protein added, which is indicated by the minus symbols. The reaction mixtures are resolved in 10 % non-denaturing polyacrylamide gels that is dried and subjected to phosphorimaging analysis

3.3.10 Rad52 mutation does not change oligomerization

ScRad52 binds to itself and forms ring-like structures *in vitro* and the self-association domain for this structure is present in the conserved N-terminus. Formation of higher order structures containing many rings have been shown for HsRad52 and depends on a C-terminal domain that has not yet been identified in ScRad52. To investigate the possibility that the four alanine substitution in Rad52-4Ala disrupts the ring structure, thus preventing repair center formation, a gel filtration chromatography analysis was performed with wild-type Rad52 and Rad52-4Ala proteins. Figure 37 shows fractions of Rad52 and Rad52-4Ala blotted with anti-Rad52 antibody. Unfortunately, both proteins are degrading somewhat in the assay and several proteolytic products are observed (figure 37). However, the alanine substitution does not appear to disrupt the oligomeric structure of the protein since Rad52 and Rad52-4Ala display identical elution profiles after gel filtration analysis.

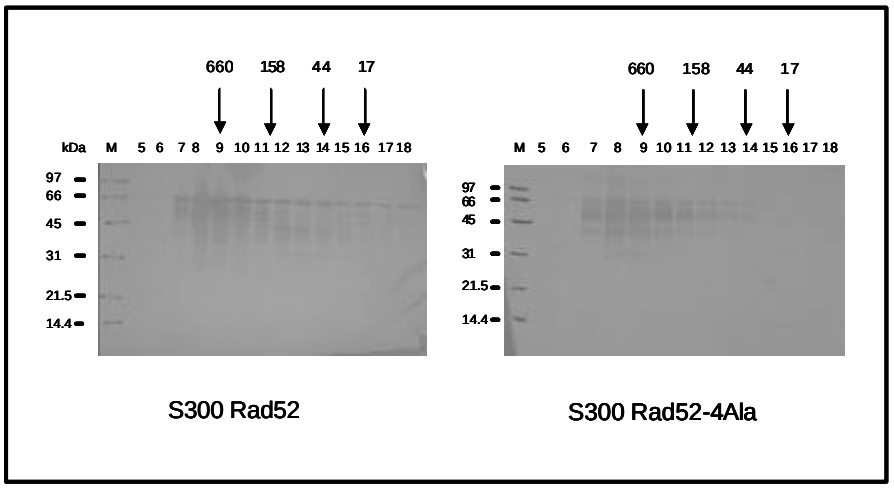


Figure 37. Rad52 and Rad52-4Ala fractions blotted with Anti-Rad52 antibody. The arrows indicate elution of a set of control proteins. The proteins are thyroglobulin (660 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa) and myoglobin (17 kDa).

3.3.11 Strand exchange

An important aspect of the Rad52 function is its role in Rad51 catalyzed strand exchange, and it was therefore tested whether the Rad52-4Ala mutant protein retains the recombination mediator activity like that of wild-type Rad52. To address this question a DNA strand exchange assay was employed to determine the recombination function of Rad51 and to test for the efficiency of the mediator activity of Rad52-4Ala. Circular single stranded DNA (css) is incubated with linear double stranded DNA (lds) to form a joint

molecule, which is converted into a nicked circular DNA duplex (nc) and linear single stranded DNA (lss) in the presence of Rad51 (figure 38A). Before testing the mediator effect of Rad52-4Ala, a Rad51 catalyzed strand exchange assay was set up followed by an assay inhibiting the strand exchange reaction with RPA.

The order of addition in the strand exchange assay is illustrated in figure 38B. Increasing concentrations of Rad51 (0 - 15 µM) was incubated with ssDNA FX174 (+) strand before addition of RPA and dsDNA. This resulted in a robust strand exchange reaction with the formation of nicked circular duplex molecules (figure 38C). Lanes 5 and 6 on figure 38C show the optimal Rad51 concentration, where the fraction of the nicked circular duplex is greatest, and thus this Rad51 concentration was used in the subsequent assays. The order of addition is critical in a strand exchange assay. If RPA is added to the ssDNA at the same time as Rad51 the strand exchange reaction is greatly inhibited with the two proteins competing for DNA binding. Next, the strand exchange assay was repeated, this time ssDNA was incubated with Rad51 and increasing amounts of RPA (4 – 7 µM) before dsDNA was added. The order of addition is shown in figure 38D. Figure 38E show the strand exchange reaction is greatly diminished with increasing RPA concentrations, lane 1 and 5 show the standard reaction control with strand exchange. Suppression of the strand exchange reaction by RPA is alleviated when the mediator Rad52 is added to the reaction. Thus, ssDNA was incubated with Rad51, RPA and increasing concentrations of Rad52 (0.6 – 1.1 µM) before dsDNA was added. Figure 38F shows the order of addition in the Rad52 mediator assay and how the strand exchange reaction runs in the presence of Rad52. Lanes 1 and 6 show the standard Rad51 reaction, lanes 2 and 7 the RPA inhibition assay and lanes 3-5 and 8-10 show the Rad52 reactions at time points 50 min. and 80 min., respectively. The assay was not optimized fully, since the inhibitory effect of RPA was not relieved completely by adding Rad52 (figure 38).

However, a slight strand exchange reaction can be detected in the strand exchange assay (figure 38. Lanes 4-5 and 9-10). The concentration of dsDNA decreases as Rad52 protein is titrated into the reaction, but it is not possible to see if there is also a nicked circular duplex formation.

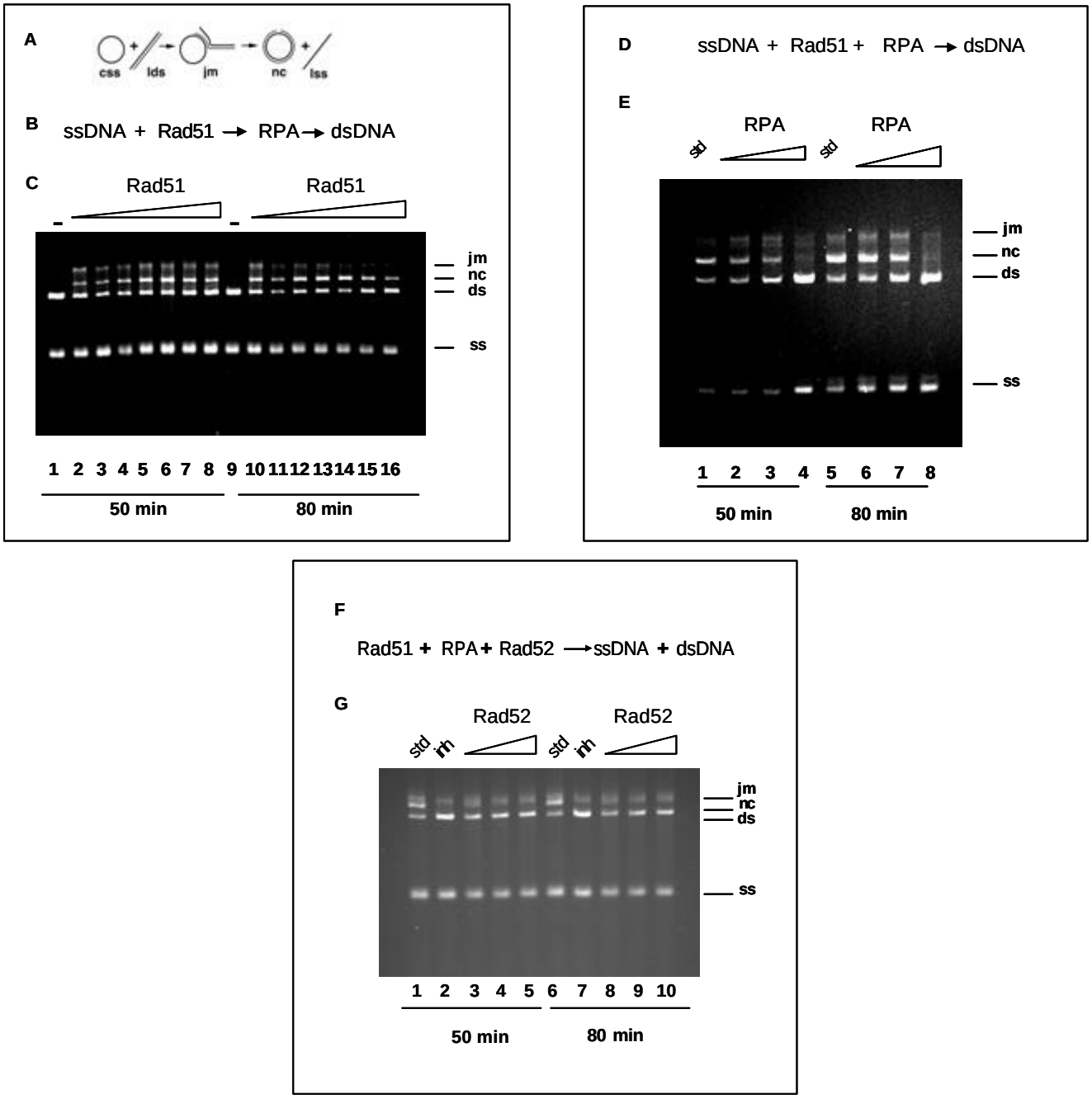


Figure 38. Strand exchange assay. A. Schematic representation of the homologous DNA pairing and strand exchange reaction. Homologous pairing between the ssDNA (ss) and linear duplex (ds) substrates yields a DNA joint molecule (jm), which is converted into a nicked circular duplex molecule (nc) by strand exchange. **B**. The FX174 ssDNA (30 µM nucleotides) is first incubated with Rad51 (0 – 15 µM) for 5 min before RPA (14 µM) is added. Following 5 minutes incubation, the linear duplex (25 µM nucleotides) is incorporated to complete the reaction. **C**. Samples are withdrawn from the Rad51 catalyzed strand exchange reaction after 50 and 80 minutes, deproteinized, and analyzed on a 0,9 % agarose gel in TAE buffer, and stained with EtBr. Samples in lanes 1 and 9 serve as control without Rad51 protein. **D**. The order of addition in the RPA inhibition assay. **E**. The reaction in panel E is the same as in panel C except that the ssDNA is

incubated both with Rad51 and RPA for 5 minutes before dsDNA is added. The concentration of RPA used is 4 µM – 7 µM. The standard Rad51 strand exchange reaction is shown in lanes 1 and 5 (std). **F**. The order of addition of proteins in the strand exchange reaction. **G**. The standard reaction (std) and RPA inhibition reaction (inh) is the same as described for C and E. In the Rad52 mediator assay ssDNA is incubated with Rad51 and RPA and increasing concentrations of Rad52 (0.6 µM – 1.1 µM). Samples are withdrawn the reaction like described in C.

Subsequently, Patrick Sung has repeated the strand exchange assay with Rad52-4Ala and found the mutant to restore strand exchange (data not shown). However, the protein preparation used was shown to be contaminated with a nuclease making it impossible to finally conclude if Rad52-4Ala indeed has a strand exchange mediator function and thus Professor Patrick Sungs laboratory is presently repeating the assay with a nuclease-free Rad52-4Ala protein.

3.3.12 Rad52 mutation does not disrupt interaction with Rad51

Considering that Rad52 acts in the strand exchange reaction by interacting physically with Rad51 and RPA it is possible that the failure of Rad52-4Ala to form repair centers is due to missing interaction between Rad52-4Ala and Rad51 or to RPA. Accordingly, pull down experiments were performed with purified Rad51 and RPA against purified Rad52 and Rad52-4Ala immobilized on nickel beads (figure 39A and B).

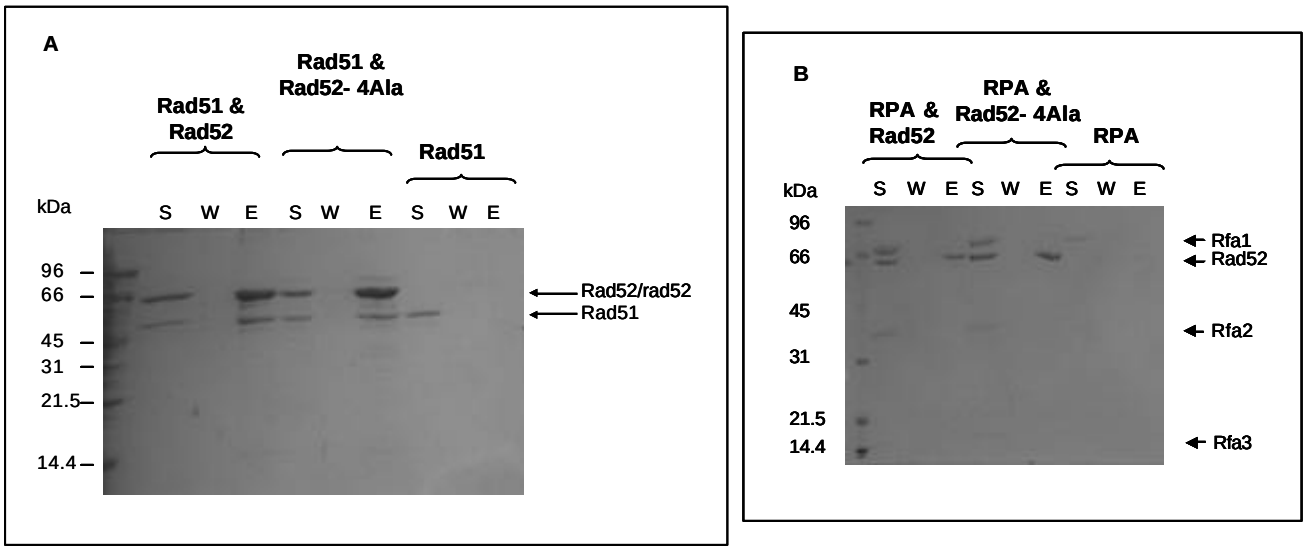


Figure 39. Pull down assays with Rad51 and RPA. Purified wild type Rad52 and Rad52–4Ala (4 µM) are mixed and preincubated with Rad51 (**A**) or RPA (**B**). Next, the protein mix and Ni-Agarose beads are mixed.

The beads are washed with buffer and then treated with SDS to elute bound Rad52 and Rad52-4Ala. The supernatant containing unbound Rad52 or Rad52-4Ala (S), the wash (W), and the SDS eluate (E) are run on 10% SDS-PAGE followed by staining with Coomassie Blue.

Figure 39A shows the pull down experiment with Rad52 and Rad52-4Ala and Rad51. Elution of Rad52 and Rad52-4Ala from the Ni-Agarose beads shows on the gel two bands. One band corresponds to the Rad52 proteins and the second band corresponds to the Rad51 protein (Figure 39, "E" lanes of Rad52 and Rad52-4Ala). It also shows both proteins present in the supernatant, but not in the wash fraction. The pull down experiment shows that Rad52-4Ala elutes Rad51 as efficiently as wild type Rad52 protein (figure 39). The negative control with Rad51 alone yields no bands in the elution showing that there is no elution of Rad51 protein from the Ni-Agarose beads themselves (figure 39A – the last lane of the gel). The result indicates a physical interaction between the purified Rad51 protein and both wild-type Rad52 and Rad52-4Ala. Figure 39B shows an attempt to pull down RPA with wild-type Rad52 and Rad52-4Ala, which unfortunately did not succeed (figure 39B). The three subunits of RPA as well as Rad52 and Rad52-4ala are present in the supernatant, but only Rad52 and Rad52-4Ala are present in the elution fraction. The assay was repeated under different conditions, but with same unsuccessful result. Lumir Krejci has since shown that Rad52 and RPA interact in the presence of DNA (data not shown), which could explain why the pull down with RPA was unsuccessful.

3.3.13 Rad52 mutation is not complemented by *RAD51* in MMS assay

The biochemical assays confirmed that Rad52-4Ala behaves like wild-type Rad52 in all *in vitro* assays tested, and did not provide an explanation for the phenotype of the mutant. Thus, the mutant strains properties were investigated in more detail *in vivo*. Since the biochemistry showed Rad52-4Ala to interact with Rad51, it was expected that overexpression of *RAD51* would suppress the MMS sensitivity and lead to higher survival of the *rad52-4Ala* strain. Thus, *RAD52*, *rad52-D327*, *rad52-4Ala* and *rad52D* strains were transformed with a galactose inducible *RAD51* overexpression plasmid (or an empty plasmid), and cell cultures treated with 0.5 % MMS for 0, 10 and 20 min. before the cell cultures were spotted on solid SC-Ura plates (data for 20 min. not shown). Cells harboring the overexpression plasmid were spotted on SC-Ura plates with galactose to induce

RAD51 transcription. The survival of *rad52-D327* was increased approx. 100-fold by *RAD51* overexpression (figure 40) whereas the *rad52-4Ala* was very sensitive and just like the *rad52D* strain did not grow following MMS-treatment. Interestingly, *rad52-4Ala* showed no increased survival following MMS treatment with and without the *RAD51* plasmid.

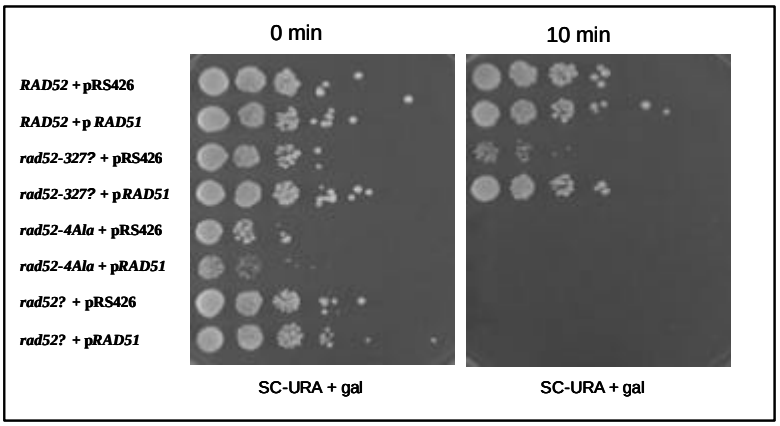


Figure 40. Overexpression of *RAD51* in *RAD52*, *rad52-D327*, *rad52-4Ala* and *rad52D* strains. The strains are transformed with *RAD51* overexpression plasmid or an empty plasmid, treated with 0.5 % for 0 or 10 min. and spotted onto SC-Ura plates with galactose. .

Not surprisingly, untreated *rad52-4Ala* cells, like *rad52D* cells, grow slowly. However, in the case of *rad52-4Ala*, this growth defect is even more pronounced when *RAD51* is overexpressed. The size of the colonies was reduced in the presence of the *RAD51* plasmid and both in the presence and absence of MMS. Thus, it was concluded that Rad51 has a dominant negative effect on the *rad52-4Ala* cells. Furthermore, *rad52-4Ala* cells were not complemented by *RAD51* overexpression unlike the other *RAD52* mutated strains. In fact, not a single colony was able to survive MMS treatment. In order to better understand this puzzling result, the *rad52-4Ala-YFP* strain was also transformed with the *RAD51* overexpression plasmid. The cells were treated with MMS and the cellular localization of Rad52-4Ala-YFP was monitored by fluorescent microscopy. In accordance with previous observations we did not identify any Rad52-4Ala-YFP foci when *RAD51* was overexpressed (data not shown). The results raise the possibility that the repair center formation of Rad52 is regulated in a more complex manner.

3.4 Discussion

In this study a mutagenesis study of *RAD52* was performed to identify regions that are important for Rad52 repair center formation and a four alanine substitution was successfully constructed, which not only abolished repair center formation, but also led to a severe MMS sensitivity.

A biochemical characterization of the Rad52-4Ala protein showed that it possesses all Rad52 functions tested, but the strand exchange assay was inconclusive and this experiment is repeated to conclude whether the Rad52-4Ala mutant protein possesses mediator activity. Unfortunately, biochemical attempts to determine an interaction with RPA were unsuccessful. Our laboratory is therefore setting up a yeast two-hybrid assay in order to determine whether RPA interacts with Rad52-4Ala. RPA binds to Rad52 through the middle region of human Rad52 and the four alanine substitution in Rad52-4Ala is located in the middle region of yeast Rad52, which makes RPA an apparent candidate in driving Rad52 into a repair focus. Alternatively, one may consider that the four alanine substitution disrupts a region that is responsible for the formation of higher order structures of Rad52. In human Rad52 this domain is located in the C-terminus like the mutation in Rad52-4Ala is. However, the results obtained from the gel filtration suggest that Rad52-4Ala is able to form higher order structures since the protein elutes the column as a huge complex. Experiment using electron microscopy would also be useful to examine this in further detail.

As mentioned earlier, Lumir Krejci performed pull down experiments with purified RPA and the middle region of Rad52, Rad52-M. Only in the presence of ssDNA he was able to detect interaction between Rad52-M and RPA. However, Hays *et al*. suggest that Rad52-RPA interaction is independent of DNA binding. They have presented Rad52 mutants that fail to interact with Rpa1 *in vivo*, but still associates with Rpa2, a subunit that does not bind DNA (Hays *et al.*, 1998). These observations show how difficult it is to mimic biological functions in the cell biochemically, and accordingly, our biochemical characterization did not provide an explanation for the phenotype of *rad52-4Ala*. The difference between the complexity in a living cell and the more simplified situation in a test tube is significant. The test tube does not include all the components found in the living cell. A repair center for

instance consists of hundreds to thousands of Rad52 molecules. Rad52 repair centers have not been reported *in vitro*, which is not surprising. It takes only a few Rad52 molecules to anneal two DNA strands (Mortensen *et al.*, 1996) and the repair centers may not be required for the *in vitro* functions of Rad52. We speculate if a more complex regulation of Rad52 repair center formation exists, which can not be picked up by biochemical assays alone. Instead, a combined study has to be conducted.

Such an attempt to better understand the assembly of repair centers has been made by Lisby *et al*. They ask the question whether the Rad52 repair centers are responsible for recruiting the DNA DSBs. If Rad52 is responsible for the assembly of several DNA DSBs in a repair focus one would expect several unassembled breaks in *rad52D* cells demonstrated by several RPA molecules bound to the processed ssDNA molecules. Experimental data using fluorescent microscopy support this idea as they observe the co-localization of two specific DNA DSBs with RPA in a *RAD52* strain, whereas RPA locates at both DNA breaks in *rad52D* cells (pers. comm. Michael Lisby).

Another argument involving Rad52 in recruiting DNA breaks to the repair center is reported by Ristic *et al.* They report that Rad52 protein complexes aggregate DNA complexes on ssDNA implying that the aggregation occurs through interaction of DNA-bound Rad52 oligomers (Ristic *et al.*, 2003). They propose that additional DNA molecules are captured by a DNA-bound Rad52 oligomer (Ristic *et al.*, 2003). The question is, is it the same mechanism that drives multiple DNA lesions into the Rad52 focus?

How does the cell benefit from repairing DNA DSBs in large repair centers? One advantage of having many repair proteins concentrated in one place could be a more rapid and efficient repair process. For example, a high concentration of Rad52 may increase the rate of Rad52 mediated reactions such as DNA annealing. The repair center is possibly a place for coordination of the different steps of the DNA repair process.

The repair center formation appears to be highly coordinated and able to process multiple lesions simultaneously. We also report some observations that indicate a complex interplay between repair proteins like Rad51 and Rad52 in response to DNA damage. The

MMS sensitivity of the truncated *rad52* alleles, *rad52-D307* and *rad52-D327* is suppressed when *RAD51* is overexpressed. The fact that the MMS sensitivity in a strain expressing a mutant Rad52 species, which lacks the entire Rad51 binding domain is fully suppressed by *RAD51* overexpression indicates that the remaining Rad52 functions are fully operational (Milne and Weaver, 1993). Accordingly, the high number of Rad52-Δ327-YFP foci likely represent stalled repair events where the turnover of a focus is impaired due to the failure to recruit Rad51, hence explaining why the number of cells containing a Rad52 focus is higher in *rad52-D327-YFP* compared to wild-type *RAD52-YFP* strain.

We propose a model for the interplay between Rad51 and Rad52 in response to DNA damage (figure 41). When a DNA DSB occur, Rad51 and Rad52 bind in small concentrations at the DNA lesion to determine if the break is a recombination substrate that needs repair by the repair center or if it for instance is a stalled replication fork that does not require repair by the HR machinery. In fact, low concentrations of Rad52 have been reported to bind DNA ends ((Van *et al.*, 1999). Fluorescent experiments show that Rad52 is at the break site before Rad51 (Lisby *et al.*, 2004), but it is possible that Rad51 arrives first yet in a low concentration, which can not be detected by fluorescent microscopy, but by ChIP analysis, which has also been suggested (Wolner *et al.*, 2003). When Rad51 and Rad52 have decided that the break is a true recombination substrate, a large concentration of Rad52 is recruited to the break site, which again recruits large concentrations of Rad51. If we accept this signalling cascade to be true, it can explain why the MMS-sensitivity of *rad52-D307-YFP* is suppressed by excess Rad51. Rad52-Δ307-YFP lacks the domain that interacts with Rad51, which makes it impossible to attract large concentrations of Rad51 to the break site. However, when *RAD51* is overexpressed, Rad51 bypasses Rad52 and gets to the lesion independently of Rad52 (figure 41).

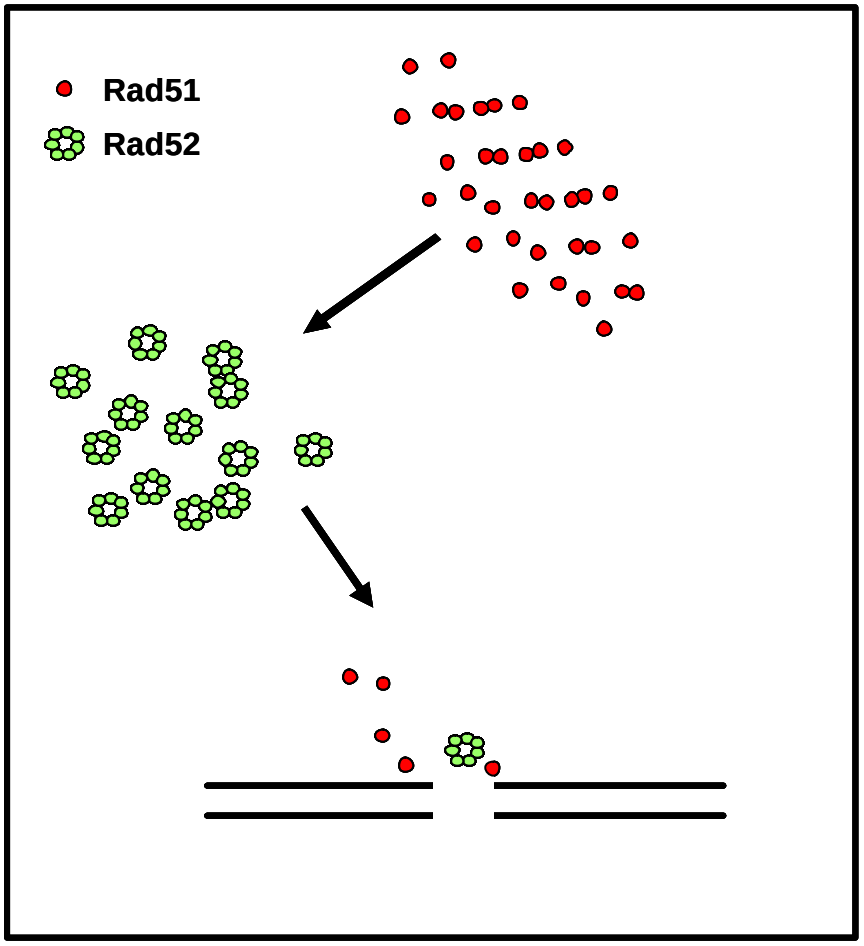


Figure 41. Model of Rad51-Rad52 signalling cascade following a DNA break. Following a DNA DSB Rad51 and Rad52 bind at the break site in low concentrations to identify recombination substrates. If the cell decides that the substrate requires recombinational repair, Rad52 is recruited to the break site in large concentrations. The elevated Rad52 concentrations at the lesion attract Rad51, which is recruited at high concentrations. Next, the lesion is repaired by the homologous recombination pathway.

It would be interesting to monitor Rad52 focus formation in these cells. More Rad52 foci would indicate that Rad51 facilitates recruitment of more Rad52 molecules to the substrate, whereas no Rad52 foci would indicate that lesions are repaired at smaller Rad52 concentrations than what can be detected by fluorescent microscopy. The next interesting question to ask is whether *rad52-D307-YFP* is able to form foci when *RAD51* is overexpressed. We know that the overexpression complements the *rad52-D307-YFP* strain and suppresses MMS sensitivity, but not if Rad51 concentrate into repair centers in this strain.

The case with the *rad52-4Ala* mutant is the complete opposite, as a *rad52-4Ala* strain is not complemented by *RAD51* overexpression. On the contrary, the MMS sensitivity of the *rad52-4Ala* strain appears to be elevated in the presence of excess Rad51. Rad52-4Ala may bind Rad51 in a deleterious way that prevents it from getting to the break site or Rad52-4Ala may block the break site physically making it inaccessible for Rad51.

In contrast to our observations, Asleson *et al*. have shown that *rad52* alleles mutated in the same region as our four alanine substitution are partly complemented by high levels of Rad51. Asleson suggests that their internal deletion creates a hole in the Rad52 protein, which makes the N- and the C-terminal come together with disastrous consequences to the cell. However, when we express a Rad52 species with internal deletion Rad52-207-237Δ-YFP-NLS it is able both to repair MMS induced lesions and to form repair centers spontaneously.

According to our model, it is possible that the four alanine substitution makes Rad52-4Ala unable to screen the DNA for lesions that require recombinational repair and initiate the signalling cascade and therewith proper repair. It is likely that the cell has a system for screening the DNA breaks to distinguish between innocent and serious lesions to prevent unnecessary recruitment of a large repair center at breaks that does not require recombinational repair. If Rad52-4Ala is unable to identify recombination substrates, but is capable of recruiting Rad51 to the break site it is possible that Rad52-4Ala and Rad51 bind innocent as well as serious lesions, which slows down the cell metabolism. This is in good agreement with our observation that overexpression of *RAD51* in *rad52-4Ala* has a dominant negative effect. In accordance with this observation, we did not detect any foci in *rad52-4Ala-YFP* when overexpressed with *RAD51*. This severe phenotype of *rad52-4Ala* could also be because it disrupts communication between functional domains in Rad52.

Another possible explanation for the Rad52-4Ala phenotype exists. The mutation in Rad52-4Ala eliminates a stretch of highly acidic amino acid residues. These acidic residues could be mimicking DNA and thereby competing with ssDNA for binding RPA. If true, Rad52-4Ala is unable to bind RPA and displace it from the ssDNA to make room for Rad51. This hypothesis is to be tested in a strand exchange mediator assay using nuclease-free Rad52-4Ala protein. This assay will show if Rad52-4Ala contains the mediator property of wild type Rad52 by displacing RPA and loading Rad51 onto the ssDNA (Jackson *et al.*, 2002) (Bochkareva *et al.*, 2005).

4. Identification of a novel DNA binding domain

Professor Patrick Sungs laboratory has recently identified a novel DNA binding domain within the C-terminal one third of ScRad52 and have used a range of biochemical *in vitro* assays to test the function of this domain in HR reactions. They report a robust DNA binding function of the C-terminus as well as DNA annealing activity and recombination mediator activity. These biochemical results propose the existence of two domains with redundant DNA binding function; one in the N-terminal and one in the C-terminal of Rad52. However, it is unknown if the C-terminal part of Rad52 has a function in HR reactions *in vivo*. To test this, a series of N-terminal Rad52 truncation mutants were constructed and analyzed in this study. Accordingly, the DNA repair function of the corresponding strains was tested in survival assays and the subcellular distribution of the mutant proteins was monitored by fluorescent microscopy. Rad52 species tagged C-terminally with fluorescent probes were unable to form repair centers spontaneously and following DNA damage. However, it was found that the C-terminal of Rad52 partly complemented the sensitivity of MMS-induced lesions in a *rad52D* strain and slightly better in a *rad52-D327* strain.

4.1 Introduction

The genetic and biochemical aspects of HR are best understood in the yeast *S. cerevisiae*, which has served as a valuable model for deciphering the intricacies of HR mechanisms. In *S. cerevisiae*, the *RAD51, RAD52, RAD54, RAD55, RAD57, RAD59*, and *RDH54* genes represent the core members of the *RAD52* epistasis group (Paques and Haber, 1999), (Sung *et al.*, 2000). In HR reactions, Rad51, the orthologue of the *Escherichia coli* RecA recombinase (Aboussekhra *et al.*, 1992), (Basile *et al.*, 1992), (Shinohara *et al.*, 1992), mediates the formation of DNA joints that link the recombining DNA molecules (Sung, 1997a). In the presence of ATP, Rad51 polymerizes onto ssDNA to form a right-handed filament, the presynaptic filament (Sung and Robberson, 1995). The presynaptic filament contains a binding site for duplex DNA. Since nucleation of Rad51 onto ssDNA is slow, presynaptic filament assembly is prone to interference by RPA. Due to its high affinity for ssDNA, RPA, if present when presynaptic filament assembly occurs, can exclude Rad51 from the ssDNA substrate and cause marked attenuation of reaction efficiency (Sung, 1994), (Sung, 1997a), (Sugiyama *et al.*, 1997).

Inclusion of Rad52 efficiently overcomes the inhibitory effect of RPA on Rad51-mediated homologous DNA pairing and strand exchange (Sung, 1997a), (Shinohara and Ogawa, 1998) an effect that can be attributed to the ability of Rad52 to nucleate Rad51 onto RPA-coated ssDNA to seed the assembly of the presynaptic filament. This replacement of RPA by Rad51 requires Rad52, providing biological relevance for the Rad52 recombination mediator activity first noted in biochemical studies.

Understanding the mechanistic basis for the Rad52 recombination mediator activity is of considerable importance, as a compromised ability to shepherd Rad51 to DNA damage sites in human cells is associated with the cancer phenotype. Specifically, there is considerable evidence that the tumor suppressor BRCA2 serves as a mediator factor in HR reactions catalyzed by human Rad51 (Jasin, 2002), (Moynahan, 2002). Molecular analyses by different groups have revealed several properties of Rad52 that are believed to be germane for its recombination mediator function. As first recognized by Mortensen *et al.* (Mortensen *et al.*, 1996) and subsequently confirmed in other studies (Shinohara *et al.*, 1998), (Van *et al.*, 1998) Rad52 binds both ssDNA and dsDNA. Rad52 assembles into oligomeric complexes that consist of seven or more protein molecules (Shen *et al.*, 1996b), (Shinohara *et al.*, 1998), (Van *et al.*, 1998), (Park *et al.*, 1996), (Stasiak *et al.*, 2000), (Ranatunga *et al.*, 2001). Rad52 also physically associates with Rad51 (Milne and Weaver, 1993), (Mortensen *et al.*, 1996), (Krejci *et al.*, 2002), as well as RPA (Park *et al.*, 1996), (Hays *et al.*, 1998), (Ranatunga *et al.*, 2001).

Ablation of the Rad51 binding domain in Rad52, as in the *rad52-?327* and *rad52-?409-412* alleles (Milne and Weaver, 1993), (Krejci *et al.*, 2002), compromises its recombination mediator function, indicating that complex formation with Rad51 as being indispensable for the recombination mediator activity of Rad52. Although not yet formally proven, it seems reasonable to propose that the DNA binding function of Rad52 is important for the effective loading of Rad51 onto ssDNA. In this regard, it has been generally assumed that the amino-terminus of Rad52, which clearly harbors a DNA binding activity, is critical for Rad52's role as recombination mediator. In addition, the oligomeric structure of Rad52 may allow it to bind DNA co-operatively, and as such, is likely also relevant for recombination mediator activity (Sung *et al.*, 2003).

In an effort to determine the molecular basis for the Rad52 recombination mediator function, the Sung laboratory identified a novel DNA binding activity within the carboxyl terminal one third of the protein. Since the C-terminal region of Rad52 also harbors the Rad51 interaction domain, they tested and verified that the carboxyl terminus of Rad52 had recombination mediator activity in isolation. Importantly, they further demonstrated that the middle portion of Rad52 binds avidly to DNA-bound RPA and makes an important contribution to the recombination mediator function. The biochemical data suggests an important function of the C-terminal of Rad52 in recombination mediator function *in vitro.* In order to determine if the C-terminal of Rad52 comprises similar activity *in vivo*, a biological characterization of the protein species was performed by MMS survival assays and fluorescent microscopy in this study. In agreement with the biochemical data, expression of the C-terminal of Rad52 partly complements MMS sensitivity in *RAD52* mutated strains. However, the C-terminal protein species were not detectable in repair centers following DNA damage. The results provide fresh insight into the complex role served by Rad52 in HR reactions.

4. 2 Materials and methods

4.2.1 Yeast strains and media

Yeast strains used in this study are displayed in table 6.

[a]Strain	Genotype
W1588-4C	MATa ade2-1 his3-11.15 leu2-3,112 ura3-1 trp1-1 LYS2 RAD52
J883	MATa ade2-1 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52-?327
UM68-5B	MATa ade2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52::HIS5
UMR101-15B	MATa ADE2 his3-11,15 leu2-3,112 ura3-1 trp1-1 LYS2 rad52::HIS5

[a]All strains are derivatives of W303-1A and W303-1B (Thomas, B. J. & Rothstein, R. 1989).
In addition to the genotype listed above all strains are RAD5

Table 6. Strains used in this study.

All media were prepared as described by Sherman (Sherman F, 1986) with minor modifications as the synthetic medium contained twice the amount of leucine (60 mg/L). All the strains used are isogenic to W303 (Thomas and Rothstein, 1989) except that they are RAD5 (Fan *et al.*, 1996), (Zou and Rothstein, 1997).

4.2.2 Construction of vectors expressing *rad52* mutant species

The CEN6 based *RAD52-YFP* expression vector, pWJ1213, and the 2 micron based *RAD52-YFP* vector pWJ1214 were used to clone *rad52-MC* and *rad52-C* and fuse with *YFP*. Inserts were constructed with PCR using pWJ1213 template and primers 169_fw_AgeI_new (5'- CTGATTGCACCGGTATGTCTTTGAGAGGGTTTGGTAA) and GFP-NLS-Stop_XhoI:

CCGTGTGCCTCGAGTCACTCGACTTTCCGCTTTTTCTTCGGAGATGCTTT GTATAGTTCATCCATGC for *rad52-MC-YFP* and 328_fw_AgeI_new (5'- CTGATTGCACCGGTATGGATCCCGTTGTAGCTAAGCA) and GFP-NLS-Stop_XhoI for *rad52-C-YFP*. Fragments were cloned into pWJ1213 and pWJ1214 using *Age*I and *Xho*I.

rad52-MC-YFP-NLS and *rad52-C-YFP-NLS* were constructed with primers 169_fw_AgeI_new and GFP-NLS-Stop_XhoI for *rad52-MC-YFP-NLS* and 328_fw_AgeI_new and GFP-NLS-Stop_XhoI for *rad52-C-YFP-NLS*. Fragments were cloned into pWJ1213 and pWJ1214 using *Age*I and *Xho*I.

Vectors expressing Rad52-MC-NLS and Rad52-C-NLS were constructed using primers 169_fw_Age_new and Rad52_NLS_stop_Nco (5' – CCGTGTGCCCATGGTCACTCGACTTTCCGCTTTTTCTTCGGAGATGCAGTAGGCTTG CGTGCATGCA) for *rad52-MC-NLS* and 328-Fw-Age_new and Rad52-NLS-stop_Nco for *rad52-C-NLS*. Fragments were inserted into CEN6 based *RAD52* expression vector pRS413 (ATCC 87518) and 2 micron based *RAD52* expression vector pRS423 (ATCC 77104) with *Age*I and *Nco*I. Figure 42 shows the *RAD52* vectors constructed for this study.

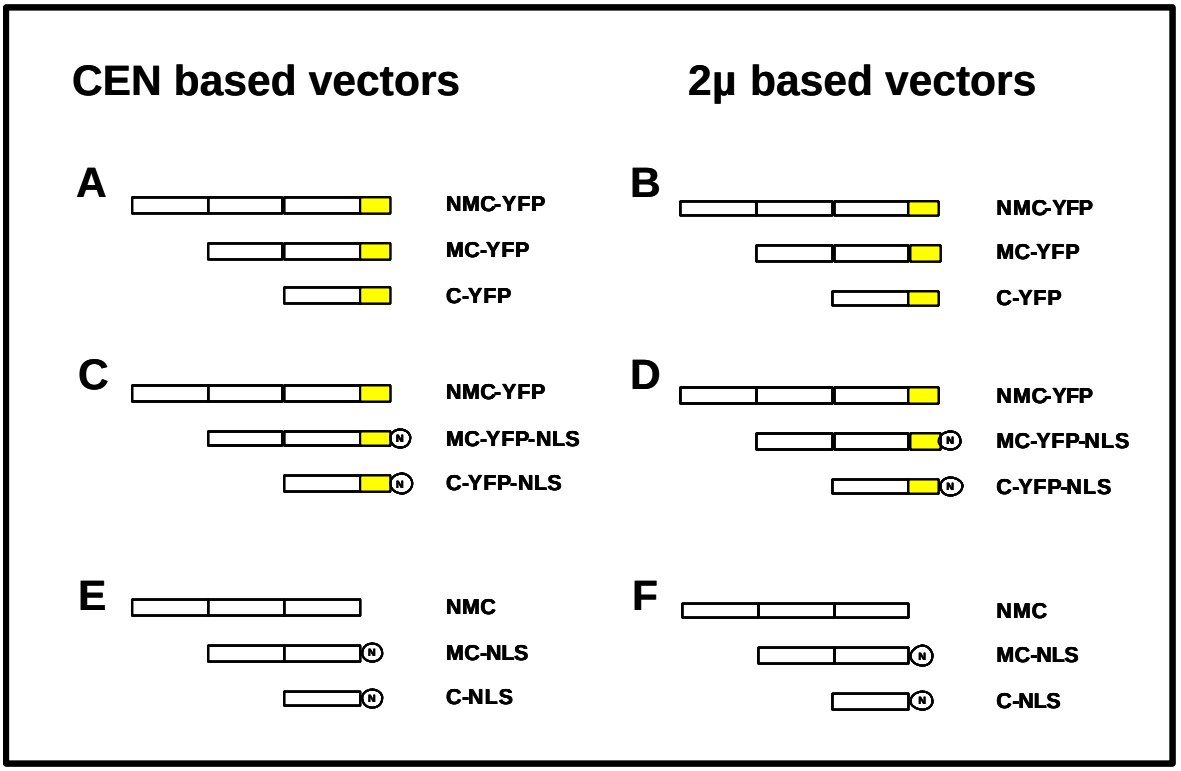


Figure 42. *RAD52* vectors used in this study. Full length Rad52 is divided into regions N, M and C. *Rad52-MC* and *rad52-C* are cloned into the CEN6 based pW1213 (*RAD52-YFP*) to be fused with *YFP* (**A**) and into the 2 micron vector pWJ1214 (*RAD52-YFP*) to be fused with *YFP* (**B**). *rad52-MC-YFP-NLS* and *rad52-C-YFP-NLS* are cloned into pWJ1213 (**C**) and pWJ1214 (**D**). *rad52-MC* and *rad52-C* are tagged with NLS and cloned into CEN6 based vector pRS413 (**E**) and pRS423 (**F**), respectively. The C-terminal N on the figure represents NLS from SV40 large T antigen.

4.2.3 Microscopy

In all experiments, cells were grown at 23°C in 3 ml overnight cultures to an OD at 600 nm of 0.3. The fluorophores used in this study is the blue- and yellow shifted enhanced variants of the *GFP* gene (Ormo *et al.*, 1996;Heim and Tsien, 1996). Cells transformed with plasmids expressing *RAD52* mutants (figure 42) were grown in SC-His medium to select for the plasmid. Repair center formation was investigated by resuspension of overnight cultures in 0.5 % MMS (M4016 from Sigma), and incubation for 15 min. Next, cultures were washed twice with water and resuspended in fresh SC-His medium. The cells were then incubated for 30 min to allow for foci to form before imaging and spun down and immobilized on a glass slide by mixing with a 37°C solution of 1.2% (wt/vol) low melting agarose (NuSieve 3:1 from FMC). Live cell images were captured with a cooled Evolution QEi monochrome digital camera (Media Cybernetics Inc., USA) mounted on a

Nikon Eclipse E1000 camera (Nikon, Japan). Images were captured at 100-fold magnification using a Plan-Fluor 100x, 1.30 NA objective lens. The illumination source was a 103W mercury arc lamp (Osram, Germany). The fluorophores YFP and CFP were visualized using a band-pass YFP filter (EX500/20, DM515, BA520 combination filter, Nikon, Japan) and a band pass CFP filter (EX436/20, DM455, BA480/40 combination filter, Nikon, Japan). Single shot image acquisition time for Rad52-YFP was 1.5 sec and 2.5 sec for CFP-Rad51 with a 12.5% neutral density (ND8) filter in place to reduce photobleaching.

4.2.4 MMS spot assay

The mutagenized plasmids were transformed into *rad52D* (UM68-5B), *rad52-D327* (J883) and *RAD52* (W1588-4C) strains. Transformed cells were grown to stationary phase in SC-His media at 30°C and resuspended in fresh SC-His media. When cell cultures reached approx. 1 x 10^8 cells/ml appropriate dilutions were made and the different cell cultures were spotted on solid SC-His plates containing 0 %, 0.0025 % and 0.005 % MMS. Plates were incubated at 30°C for 2 days before examination (Prakash and Prakash, 1977a).

4.3 Results

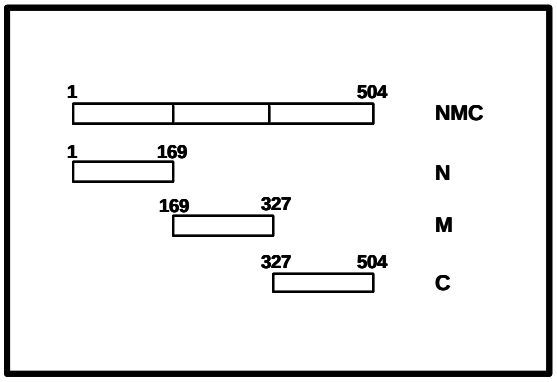


Figure 43. Schematic representation of Rad52 constructs. Full length Rad52 is divided into regions N, M and C. Region N expresses aa 1 - 169, region M aa 169 – 327 and region C aa 327 – 504.

4.3.1 *In vitro* characterization of novel DNA binding domain in C-terminal of Rad52 (performed by Professor Patrick Sungs group - data not shown)

In Professor Patrick Sungs laboratory, GST-tagged (N-terminal) forms of full length Rad52 and variants that encompass the N-terminal, the middle and the C-terminal (figure 43) and

a combination of these were purified. The protein species were analyzed by biochemical *in vitro* assays to determine their function in HR reactions. Using DNA binding assays it was shown that both Rad52-N and Rad52-C harbors DNA binding function. In fact, Rad52-C seemed more adept at binding both ssDNA and dsDNA than the known DNA binding domain in Rad52-N. Furthermore, DNA annealing experiments showed that Rad52-C, like full length Rad52 and Rad52-N, exhibited DNA annealing activity. Rad52-C had a higher annealing efficiency than Rad52-N. Moreover, Rad52-C was capable of annealing RPA-coated DNA strands, again more efficiently than Rad52-N.

Since the Rad51 binding domain is located in the C-terminus of Rad52, they addressed whether Rad52-C has recombination mediator activity. Using a DNA strand exchange assay (for details on the assay see section 3.2.11), they demonstrated that Rad52-C restores a significant level of DNA strand exchange with RPA co-addition, although, when compared to full length Rad52, higher amounts of Rad52-C protein is needed and the final extent of restoration is significantly lower than what could be achieved with full length Rad52. Either Rad52-N or Rad52-M is capable of DNA strand exchange restoration. Since full length Rad52 was considerably more effective than Rad52-C in the DNA strand exchange restoration, they investigated whether the N- and M- portions contribute to the mediator function of the full length protein. Thus, they fused either the N- or M-portion to the C-terminus and repeated the strand exchange assay, and found Rad52-MC to be significantly more effective than Rad52-C in promoting the restoration of DNA strand exchange. The M-portion of Rad52 has no known DNA binding activity and no other defined function required in the recombination mediator activity. However, in human Rad52 the M-portion interacts with RPA. In order to test if Rad52-M increased the mediator function of Rad52-C through interaction with RPA they performed a pull down assay (for details on the assay see section 3.2.9) with Rad52-M and RPA bound to ssDNA. The pull down showed Rad52-M to associate with the DNA-bound form of RPA.

4.3.2 Low expression of *rad52-MC-NLS* and *rad52-C-NLS* does not complement MMS sensitivity

To test if the C-terminal DNA binding domain of Rad52 has any biological relevance, the Rad52-MC and Rad52-C mutant proteins were tested *in vivo*. In order to determine the

effect of the C-terminal *rad52* mutants in HR, a survival assay was performed to test if the mutant proteins were able to suppress MMS sensitivity. However, in chapter 2 it was shown how Rad52 mutant proteins devoid of the N-terminal of Rad52 are unable to self-associate and therefore mis-sort. Instead of entering the nucleus they remain in the cytosol. Rad52-MC and Rad52-C both lack the N-terminal part of the Rad52 sequence and it is reasonable to expect them to be unable to sort correctly to the nucleus. In agreement with these assumptions, the fusion proteins Rad52-MC-YFP and Rad52-C-YFP both were found to locate in the cytosol when investigated by using fluorescent microscopy (figure 46). Consequently, Rad52-MC and Rad52-C were tagged with NLS from SV40 virus to ensure nuclear localization (figure 47).

Next, *rad52-MC-NLS* and *rad52-C-NLS* were expressed from CEN6 based vectors in *rad52D*, *rad52-D327*, and *RAD52* strains. The *rad52-D327* strain was included in this study since the ablation of the Rad51 binding domain compromises the recombination mediator function of Rad52. The *RAD52* strain was chosen as control to test for dominant negative effects, but none of the constructs had that effect (data not shown). The transformants were tested for their ability to perform Rad52 functions during repair of MMS induced DNA damage in a survival assay. The cell cultures were spotted on MMS-containing Sc-His plates in serial 10-fold dilutions (figure 44). At the dose used, no significant complementation of either of the strains was observed. The first row on figure 44 shows *rad52D* cells expressing wild type *RAD52* spotted on plates with increasing concentrations of MMS. Wild type *RAD52* complements the MMS sensitivity since *rad52D* cells expressing *RAD52* are only mildly affected even at 0.005 % MMS. The next two rows illustrate *rad52D* cells expressing *rad52-MC-NLS* and *rad52-C-NLS*. The last row shows cells transformed with an empty vector. There is clearly no significant increase in the survival of cells expressing any of the two mutant proteins and cells with the empty vector. This result could be because the low copy CEN vector is maintained at few copies per cell resulting in Rad52-MC-NLS and Rad52-C-NLS concentrations that are too low to suppress the MMS sensitivity. The lower panel shows a similar experiment in *rad52-D327* cells and even though the cells are not as sensitive towards MMS treatment as *rad52D,* but the same result is obtained with no complementation by the mutant proteins.

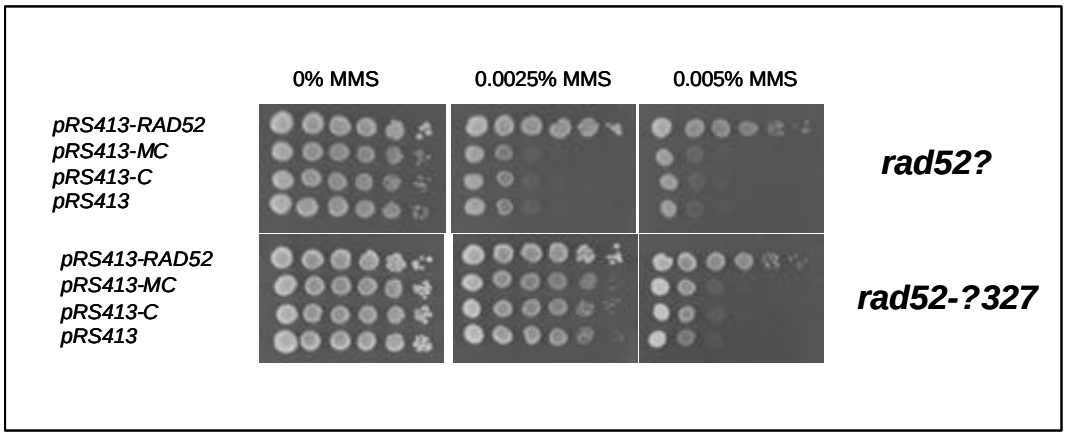


Figure 44. MMS spot assay shows no complementation when *rad52-MC-NLS* and *rad52-C-NLS* are expressed from low copy plasmids.

rad52D and *rad52-D327* strains are transformed with pRS413-*RAD52*, pRS413-rad52-*MC-NLS*, pRS413-*rad52-C-NLS* and pRS413. Cell cultures, 10^8 cells/ml, are diluted in 10-fold series and spotted onto SC-His plates containing 0, 0.0025 and 0.005 % MMS. Pictures are captured after two days incubation.

4.3.3 High expression of *rad52-MC-NLS* and *rad52-C-NLS* complements MMS sensitivity

Next, *rad52-MC-NLS* and *rad52-C-NLS* were constructed and expressed from a high copy 2 micron plasmid in *rad52D* and *rad52-D327* strains. The transformants were subsequently tested in a survival assay like the one described above. Results from the assay showed a slight complementation of the *rad52D* strain in the presence of Rad52-MC-NLS and Rad52-C-NLS proteins (figure 45). When *rad52D* cells harboring an empty vector are compared with cells expressing *rad52-MC-NLS* and *rad52-C-NLS* there is a 10-fold increase in the survival of the cells expressing the mutant proteins (top panel – figure 45). A more robust complementation is observed in the *rad52-D327* strain, which displayed a 10 – 100 fold increased survival when comparing strains expressing *rad52-MC-NLS* and *rad52-C-NLS* to strains harboring the empty vector (bottom panel - figure 45). The result suggests that the endogenously expressed Rad52-Δ327 protein and the mutant Rad52-MC-NLS and Rad52-C-NLS can collaborate in the repair process when the latter proteins are expressed at high copies. When a *RAD52* strain was transformed with 2 micron plasmids expressing the *rad52* mutant alleles *rad52-MC-NLS* and *rad52-C-NLS* no change in resistance to MMS was observed compared with strains transformed with an empty vector (data not shown). This suggests that there is no dominant negative effect of the two mutagenized vectors. Dornfeld and Livingston have tried to suppress MMS sensitivity in a

rad52 deficient strain by overexpressing *RAD52* under the control of the constitutive promoter Eno1 (Dornfeld and Livingston, 1991). Overexpression of *RAD52* was tested both in a *RAD52* strain and a *rad52* deficient strain, and the MMS resistance was not increased in the *RAD52* strain and increased to wild type levels in the *rad52* deficient strain. This shows that Rad52 is not rate limiting for DNA damage repair in wild type *RAD52* strains, which is in good agreement with the results obtained (in this study) for wild type *RAD52* strains (data not shown).

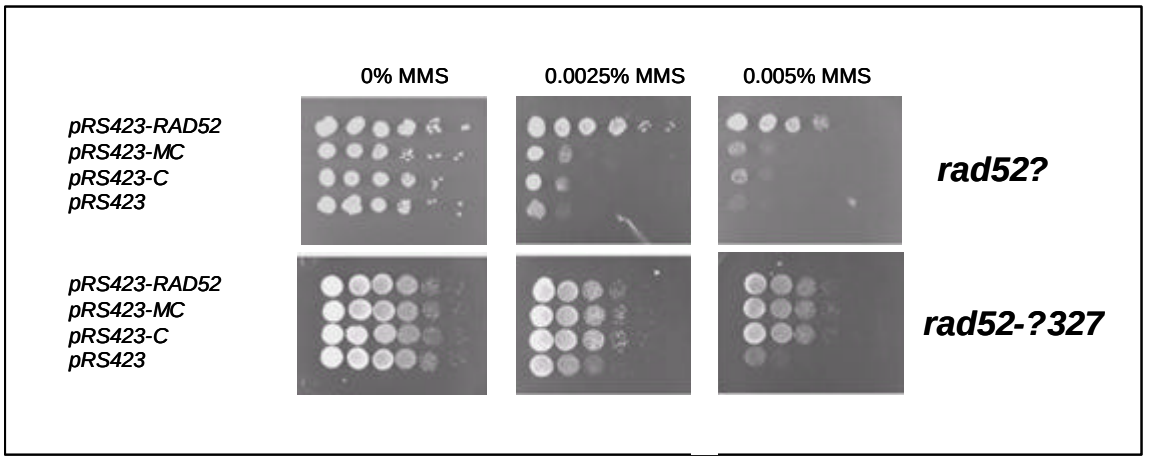


Figure 45. MMS spot assay show complementation of a *rad52-D327* strain when *rad52-MC-NLS* and *rad52-C-NLS* are expressed from high copy plasmids. *rad52D* and *rad52-D327* strains are transformed with pRS423-*RAD52*, pRS423-*rad52-MC-NLS*, pRS423-*rad52-C-NLS* and pRS423. Cell cultures, 10^8 cells/ml, are diluted in 10-fold series and spotted onto SC-His plates containing 0, 0.0025 and 0.005 % MMS. Pictures are captured after two days incubation.

4.3.4 Rad52-MC-YFP and Rad52-C-YFP locates in the cytosol

The survival assay above showed a biological relevance of the C-terminal DNA binding domain, yet it is unknown what role this novel domain has in the repair process. Thus, the protein species were investigated by using fluorescent microscopy. Initially, the C-termini of Rad52-MC and Rad52-C were extended by YFP to allow detection of the cellular distribution of the resulting fusion proteins. The resulting vectors harboring *rad52-MC-YFP, rad52-C-YFP* and *RAD52-YFP* were expressed in a *rad52D* strain. *rad52D* cells expressing wild type and the mutated Rad52 proteins were analyzed by using fluorescent microscopy, which showed both Rad52-MC-YFP and Rad52-C-YFP to locate in the cytosol, whereas wild type Rad52-YFP located in the nucleus (figure 46). As mentioned above, this result is

in accordance with the model proposed in chapter 2 that Rad52 is transported to the nucleus only if it contains NLS 2 as well as the N-terminal self-association domain. Both Rad52-MC and Rad52-C lack the self-association domain and only Rad52-MC contains the NLS 2 sequence, which could explain the mis-sorting. Gel filtration assays with purified Rad52-N, Rad52-M and Rad52-C conducted by the Sung laboratory back this up (data not shown). They discovered an oligomeric nature of Rad52-N and a monomeric nature of Rad52-M and Rad52-C. Since the N-terminal is a prerequisite for association of the Rad52 proteins and nuclear transport it is expected that Rad52-MC and Rad52-C also are monomeric protein species *in vivo*.

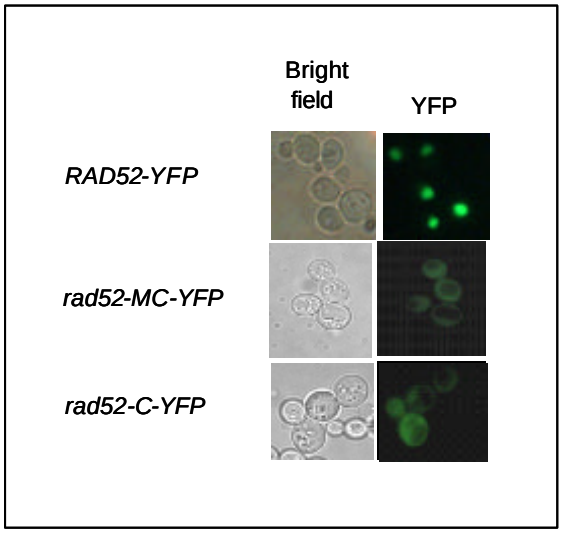


Figure 46. Cellular distribution of Rad52-YFP, Rad52-MC-YFP and Rad52-C-YFP. Rad52-MC-YFP and Rad52-C-YFP both locate in the cytosol when expressed in a *rad52D* strain whereas wild type Rad52-YFP is located in the nucleus. Pictures are pseudocolored using Adobe Photoshop software.

4.3.5 Rad52-MC-YFP-NLS and Rad52-C-YFP-NLS locate in the nucleus

To solve the problem of the above, a new set of NLS-tagged *rad52* alleles were constructed. Accordingly, plasmids were constructed to express *rad52-MC-YFP* and *rad52-C-YFP* tagged C-terminally with the prototypic SV40 large T antigen NLS and transformed into a *rad52D* strain to monitor the proteins localization by using fluorescent microscopy. The NLS-tagging resulted in a translocation of the protein species to the nucleus (figure 47).

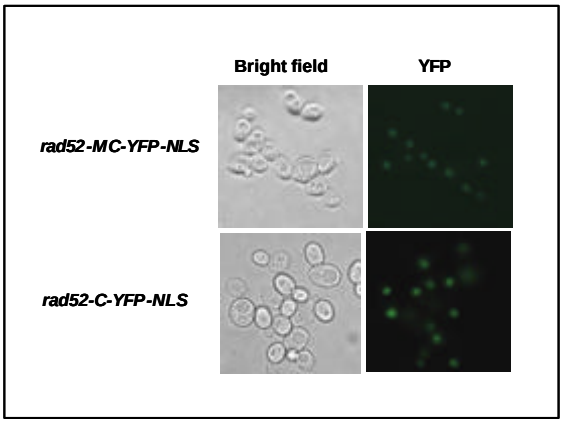


Figure 47. Cellular distribution of Rad52-MC-YFP-NLS and Rad52-C-YFP-NLS. Both proteins locate in the cytosol when expressed in a *rad52D* strain. Pictures are pseudocolored in Adobe Photoshop.

4.3.6 Rad52-MC-YFP-NLS and Rad52-C-YFP-NLS fail to form repair centers in response to DNA damage

In this thesis, it has been discussed how Rad52-YFP forms repair centers in response to DNA damage, and that these bright foci represent ongoing repair in the cell. Accordingly, it was investigated if Rad52-MC-YFP-NLS and Rad52-C-YFP-NLS also have the capability to form repair centers when treated with the alkylating agent MMS. In chapter 3 it was demonstrated how a region in the Rad52 sequence is required for the recruitment of Rad52 to a repair center. The sequence in Rad52 that was found to be important includes amino acid residues 299-311 and these particular amino acid residues are present in both the Rad52-MC-NLS and Rad52-C-NLS mutant proteins. For that reason, it was tested if the truncated protein fusions accumulate in response to DNA damage like wild type Rad52-YFP does. CEN based and 2 micron based vectors harboring *rad52-MC-YFP-NLS* and *rad52-C-YFP-NLS* were transformed into a *rad52D* strain and transformants were treated with MMS and analyzed by fluorescent microscopy at times 0, 30 and 60 minutes. All focal planes of the cells were inspected for Rad52-MC-YFP-NLS and Rad52-C-YFP-NLS foci, but no foci were observed in any of the cell cultures when a population of more than 100 cells was inspected for each time point (data not shown). The reason that no foci were observed could be because it takes more than 60 minutes for the mutant proteins to locate to the break site, or because the mutant proteins only rarely form repair centers or that the mutant proteins function in concentration that are too low for detection by fluorescent microscopy.

Since Rad52-MC-NLS and Rad52-C-NLS to a small degree complement MMS sensitivity in a *rad52D*, it suggests a biological role of the mutant proteins in HR. The complementation was significantly better when the mutant proteins were expressed in a *rad52-D327* strain and it is therefore tempting to speculate that the two “halves” of Rad52, one part expressed from the plasmid and one endogenously come together and cooperate to repair the DNA damage. This idea was tested by using fluorescent microscopy. A *rad52-D327-CFP* strain was transformed with 2 micron based vectors carrying *RAD52-YFP*, *rad52-MC-YFP-NLS* and *rad52-C-YFP-NLS* and the transformants were inspected like described above. The endogenously expressed Rad52-Δ327-CFP formed foci in response to the MMS treatment and so did wild type Rad52-YFP. However, none of the

truncated proteins Rad52-MC-YFP-NLS or Rad52-C-YFP-NLS were observed to form foci (figure 48).

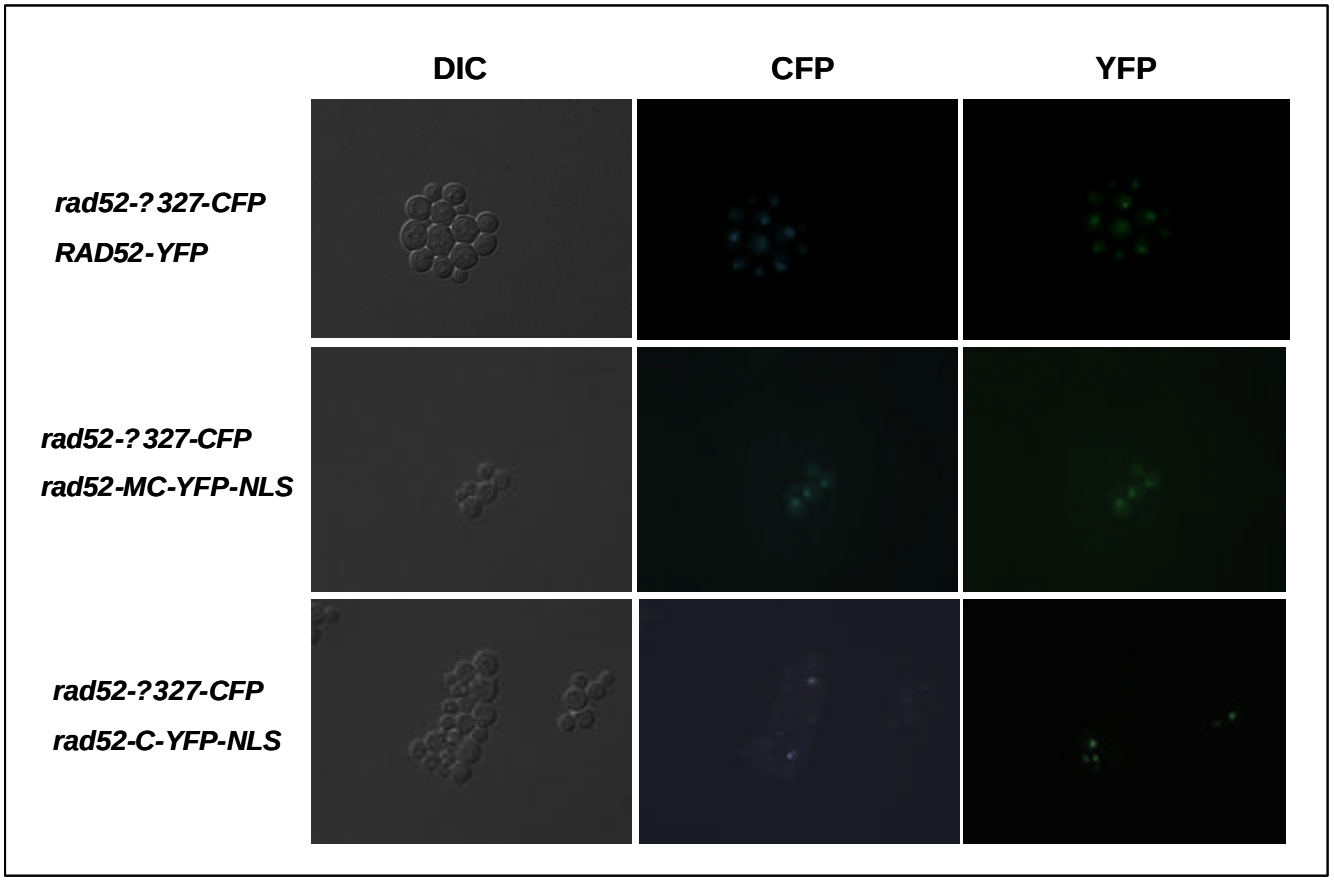


Figure 48. Focus formation of Rad52-Δ327-CFP, Rad52-YFP, Rad52-MC-YFP-NLS and Rad52-C-YFP-NLS following MMS treatment. Rad52-Δ327-CFP forms foci in response to DNA damage. Rad52-YFP does also form foci, which co-localize with the Rad52-Δ327-CFP foci. Rad52-MC-YFP-NLS and Rad52-C-YFP-NLS do not form foci.

The result indicates that the C-terminal portion of Rad52 found in Rad52-MC-YFP-NLS and Rad52-C-YFP-NLS do not concentrate at the break site in detectable concentrations.

4.3.7 Co-expression of *rad52-Δ327-YFP* and *rad52-C-NLS* does not make CFP-Rad51 form repair centers

Lastly, Rad52's role in repair center formation of Rad51 was investigated. The Rad51 protein has been shown to concentrate into repair centers in a Rad52-dependent manner (Lisby *et al.*, 2004). However, the concentration of Rad51 into repair centers appears to be

mediated through a physical interaction with Rad52 since Rad51 is unable to form repair center in a *rad52-D327* strain background (PhD-thesis, Tanja Thybo). To test if the addition of the Rad51 interaction domain found in Rad52-C-NLS can facilitate Rad51 repair center formation either wild type *RAD52* or *rad52-C-NLS* were expressed in a *rad52-D327 CFP-RAD51* strain background the sub-cellular distribution of the CFP-Rad51 protein was monitored. The top panel in figure 49 shows how CFP-Rad51 repair centers are formed in the presence of wild type Rad52. The lower panel shows the cellular distribution of CFP-Rad51 when endogenous *rad52-D327* and *rad52-C-NLS* from a plasmid are co-expressed (figure 49) with no CFP-Rad51 repair centers formed.

.

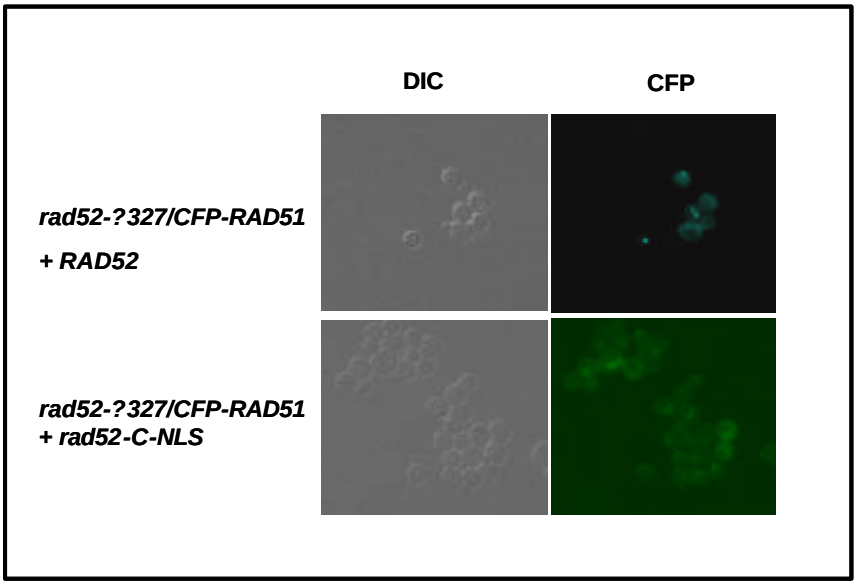


Figure 49. Co-expression of Rad52 truncations *rad52-D327-YFP*, *rad52-C-NLS* together with *CFP-RAD51*. After MMS treatment, co-expression of wild type *RAD52* and *CFP-RAD51* leads to focus formation of CFP-Rad51. Expression of two halves of *RAD52*; *rad52-D327-YFP* and *rad52-C-NLS* does not promote CFP-Rad51 repair center formation. Pictures are pseudocolored in Adobe Photoshop.

The result indicates that the two Rad52 protein species expressing the N-terminal (Rad52-Δ327) or the C-terminal (Rad52-327Δ) are unable to collaborate efficiently to facilitate the repair center formation of CFP-Rad51 when they are co-expressed. The Rad52-MC-YFP-NLS and Rad52-C-YFP-NLS proteins were also unable to form repair centers in response to DNA damage, which indicates that the mutant proteins have reduced function in the DNA repair process. One could speculate that the C-terminal of Rad52 is insufficient for efficient break site assembly and thus when the two halves of the proteins are co-expressed they are physically separated in the cell. Rad52-Δ327 is recruited to the break

site whereas Rad52-C-NLS is not. Taken together these results suggests separate roles of the N- and the C-terminal DNA binding domains of Rad52 in the DNA DSBR pathway.

4.4 Discussion

The result of this work show that the C-terminal DNA binding domain recently identified in Professor Patrick Sungs laboratory also is biological functional. We observe that MMS sensitivity is complemented in a *rad52-D327* strain when the C-terminal part of the Rad52 protein, Rad52-MC-NLS or Rad52-C-NLS, is present. However, the fluorescently tagged counterparts of these truncated proteins can not form repair centers, in response to DNA damage in a *rad52-D327* strain, that are visible by fluorescent microscopy. In addition, the two halves of the Rad52 protein are not able to collaborate efficiently to mediate Rad51 repair center formation. There is some collaboration between the two halves of the Rad52 protein in the survival assay, but it seems that the split protein does not come together in concentrations that are possible to detect by fluorescent microscopy.

But what is the biological function of the additional DNA binding domain in the C-terminal, and why does the RAd52 protein have two DNA binding domains? The two DNA binding domains may not be completely redundant, but could play different roles in the homologous recombination process. One could imagine that the two domains have some overlapping function as well as unique function. For instance, Rad52 protein species expressing only the N-terminal DNA domain like Rad52-Δ327-YFP binds at the break site and concentrate into repair centers in response to DNA damage, but the cells are sick and sensitive to DNA damage. Rad52-Δ327 may be sufficient to bind at the lesion, but Rad51 is needed to complete the repair process and Rad52-Δ327 is unable to attract Rad51 to the site of damage. When the remaining part of the Rad52 protein in the form of Rad52-MC-NLS or Rad52-C-NLS is added to the reaction we observed that the *rad52-D327* cells were less sensitive to MMS induced damage. One explanation for this might be that the C-terminal part of Rad52 is attracting Rad51 to the break site. Rad52-MC and Rad52-C might form an intermediate with the ssDNA via the DNA binding domain in the C-terminal. Next, Rad51 might be attracted to the complex allowing the repair process to continue. A mediator role depending on this region of Rad52 is backed up by the *in vitro* results that showed that the C-terminal interacts with both Rad51 and RPA.

In contrast, when only the C-terminal DNA binding domain is present like in Rad52-C-YFP-NLS, the protein does not concentrate at the break site in concentrations that are

detectable by fluorescent microscopy. The explanation for this could be that the presence of the N-terminal DNA binding domain is necessary for efficient recognition of the homologous recombination substrate before the C-terminal of Rad52 binds to the ssDNA or it could be that Rad52-C-YFP-NLS finds the lesion, but only binds in concentrations that are too low for detection.

In addition, the collaboration with Rad51 appears to be different between Rad52 species expressing the N-terminal DNA binding domain and species expressing the C-terminal domain. Overexpression of *RAD51* can suppress the MMS sensitivity of a *rad52-D327* strain, which suggests a functional interaction between the two proteins. In this situation, there is no Rad51 binding domain present in Rad52 and Rad51 might be able to be recruited to the break site by simple diffusion. When a Rad51 binding domain is present, like in Rad52-MC-NLS or Rad52-C-NLS, the two proteins interact physically, but do not concentrate at the break site in detectable levels. This suggests that the N-terminal DNA binding domain is responsible for binding to the DNA break site.

5. Discussion and conclusion

In this study several functional domains of Rad52 have been investigated both *in vitro* and *in vivo*. It has been shown how different methods such as biochemical assays, genetics and fluorescent microscopy can be used together to investigate the Rad52 proteins role in homologous recombination. Thus, it was shown that fluorescent microscopy is a valuable tool in the investigation of DNA repair by homologous recombination. Not only was it possible to determine the sub-cellular localisation of DNA repair proteins in yeast cells, but also to monitor the translocation of these in response to DNA damage. In addition, the technique proved to be useful in the functional characterisation of a DNA binding domain as wells as protein interaction between mutated repair proteins *in vivo*. The functional domains of Rad52 described in this thesis were successfully identified by the use of genetic engineering of *RAD52* strains. Rad52 alleles were truncated, mutated and tagged with fluorescent probes in order to determine regions involved in both nuclear transport and repair center formation. The ability of a mutant protein to repair DNA DSBs was monitored in survival assays by using MMS to introduce the breaks. It would also be interesting to investigate these mutant's DNA repair properties in homologous recombination assays.

Biochemical assays were used to try and explain what causes the phenotype of a repair center formation deficient Rad52 mutant protein. The mutant protein was able to perform many known Rad52 functions *in vitro.* Still, it has not been established whether the mutant protein has mediator function. Further investigation of the collaboration between Rad51 and Rad52 and between Rad52 and RPA in this repair center formation is necessary to fully understand the mechanism by which the proteins are recruited and subsequently concentrated at the break site.

The results obtained in this work lead to the conclusion that an NLS (NLS 2) in the Rad52 protein sequence is necessary for nuclear transport of Rad52 and not for other Rad52 functions. It was also concluded that this sequence alone is not enough to ensure nuclear localization, but that the Rad52 protein also has to multimerize to be transported to the nucleus. Next, it was concluded that an acidic region in the Rad52 protein is necessary for

repair center formation as well as for repair of MMS induced DNA DSBs. We suggested a model where Rad51 and Rad52 bind at the break site in low concentrations to determine if the break is a homologous recombination substrate that requires repair by the homologous recombination machinery. If that is the case more Rad52 is recruited to the break site, which again attracts more Rad51 and the lesion is repaired. Rad52-4Ala still bind DNA and Rad51, but it could be unable to distinguish between lesions that require the DNA repair machinery and which that do not, which can have catastrophic consequences to the cell if repair is initiated repeatedly whenever a break occurs. Another possible explanation presented is that the mutation in Rad52-4Ala eliminates an acidic stretch of amino acid residues. This stretch mimics DNA and makes it possible for Rad52 to compete with ssDNA for RPA binding. Without this property, Rad52-4Ala is unable to bind RPA and load Rad51 onto the ssDNA. Finally, survival assays were used to show that the C-terminal DNA binding domain functions *in vivo* although it is unable to form repair centers. Thus, we speculate if there are separate roles for the redundant N- and C-terminal DNA binding domains in the homologous recombination pathway. The first DNA binding domain appears to be necessary for Rad52 binding at the break site, whereas the DNA binding domain in the C-terminal seems unable to recruit Rad52 to the break site. However, the C-terminal DNA binding domain may be responsible for forming an intermediate with ssDNA and subsequently attracting Rad51 to the break site.

.

Rad52 is a multifunctional protein, which has been further demonstrated in this thesis by the identification of novel functional domains. The presence of multifunctional proteins in yeast cells could be due to the fact that the yeast genome is much smaller than the human genome. With fewer proteins in yeast cells it is an advantage to the cell with multifunctional proteins. Through evolution these functions may have split into several proteins in higher eukaryotes. Higher eukaryotes have more complex systems than yeast. For instance *S. cerevisiae* has only one known Rad51 whereas mammalian cells have a family of Rad51 proteins. The same rationale could be used to explain the different role that Rad52 plays in yeast and in mammalian cells. For instance, human *RAD52* and *XRCC3* are synthetic lethal implicating redundant function of the two proteins in the DNA DSBR pathway (Fujimori *et al.*, 2001). It is tempting to speculate that protein functions of yeast Rad52 is maintained by mammalian proteins like BRCA2 or XRCC3. In yeast, a deletion of *rad52*

has severe consequences, whereas the effect of such a disruption has less impact in mammalian cells. When *rad52* is disrupted in yeast it means that several functions are impaired and consequently the cell is unable to survive DNA damage. In mammals however, the impact of disrupting *rad52* is less severe probably because the "Rad52-functions" are maintained by more than one protein.

The results presented in this thesis results in a new functional map of the Rad52 protein, which is presented below (figure 50).

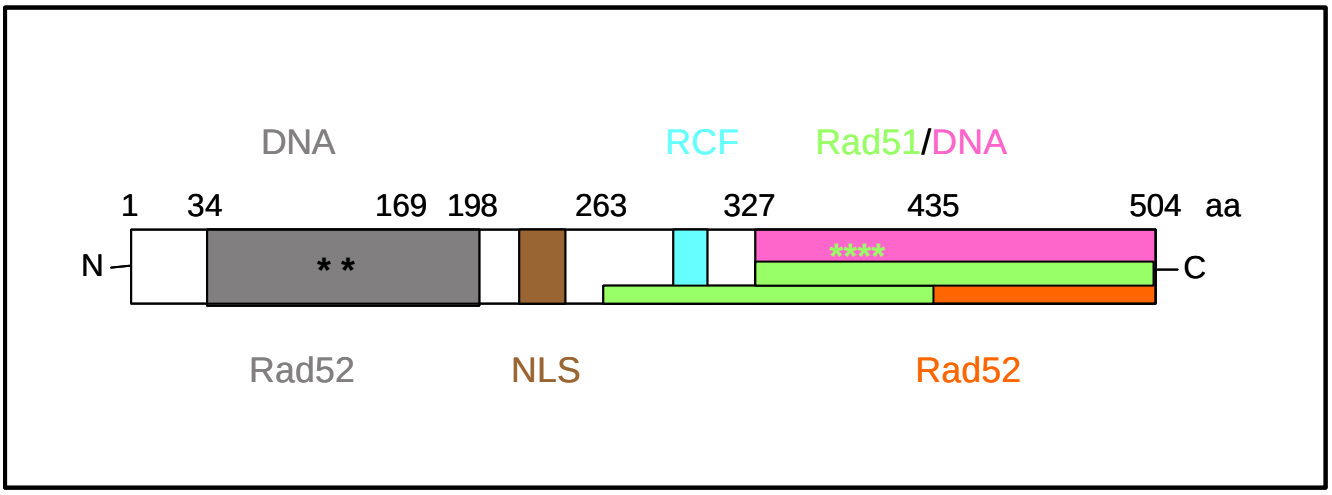


Figure 50. Map of *S. cerevisiae* Rad52 with new functional domains on. The outline shows the 504 aa of ScRad52. The grey region spanning aa 34 – 198 corresponds to the evolutionary conserved region of ScRad52 (Mortensen *et al.*, 1996). DNA binding is mediated through the N-terminal, which is also responsible for HsRad52 self-association (Shinohara *et al.*, 1998), (Ranatunga *et al.*, 2001). A Rad51 interaction domain is located in the C-terminus. The upper green section depict the Rad51 interaction domain identified by Milne and Weaver (Milne and Weaver, 1993) and the lower green section the domain Mortensen *et al.* found (Mortensen *et al.*, 1996). The C-terminal asterisks mark the region including aa 409 - 412 identified as responsible for Rad51 interaction by Krejci and co-workers (Krejci *et al.*, 2002). Amino acid substitutions G121E and G142D in ScRAD52 disrupting the interaction to Rfa1 (Hays *et al.*, 1998) are shown by asterisks in the N-terminus. HsRad52 binds to itself through the N-terminus (Shen *et al.*, 1996b) as well as the C-terminus (Ranatunga *et al.*, 2001), and binds to RPA through the middle region (Park *et al.*, 1996), (Ranatunga *et al.*, 2001). The NLS sequence responsible for Rad52 nuclear sorting is shown in brown and is spanning aa residues 231 – 237. In blue is shown the aa residues responsible for Rad52 repair center formation. These residues include 308 – 311. In red is shown the newly identified DNA binding domain located in the last third of the Rad52 protein sequence.

6. Appendix

6.1 Rad52 sequence alignment

The alignment between amino acid sequences of Rad52 from yeast, mouse and human.

```
S.cerevi              MAFLSYFATENQQMQTRRLPRTAEGSGGFGVLLMNEIMDMDEKKPVFGNHSEDIQTKLDK 60
K.lactis              ------------MEDT-------GSG--------------------KNGKDDIQTKLDK 20
H.sapien              -----MSGTEEAILGGRDSHP-AAGGGSVLCFGQCQ---------YTAEEYQAIQKALRQ 45
M.muscul              -----MAGPEEAVHRGCDNHPPFVGGKSVLLFGQSQ---------YTADEYQAIQKALRQ 46
                            ....             *. .       .           :  : **. * :

S. cerevi              KLGPEYISKRVGFGTSRIAYIEGWRVINLANQIFGYNGWSTEVKSVVIDFLDERQGKFSI 120
K. lactis              KLGPEYISKRVGFGSSRVAYIEGWKAINLANQIFGYDGWSTEVKNVTIDFLDERQGRFSI 80
H. sapien              RLGPEYISSRMAGGGQKVCYIEGHRVINLANEMFGYNGWAHSITQQNVDFVDLNNGKFYV 105
M. muscul              RLGPEYISSRMAGGGQKVCYIEGHRVINLANEMFGYNGWAHSITQQNVDFVDLNNGKFYV 106
                       :*******.*:. * .::.**** :.*****::***:**: .:..  :**:* .:*:* :

S. cerevi              GCTAIVRVTLTSGTYREDIGYGTVENERRKPAAFERAKKSAVTDALKRSLRGFGNALGNC 180
K. lactis              GCTAIVRVSLADGTFREDIGYGTVENERRKASAFERAKKSAVTDALKRSLRGFGNALGNC 140
H. sapien              GVCAFVRVQLKDGSYHEDVGYGVSEGLKSKALSLEKARKEAVTDGLKRALRSFGNALGNC 165
M. muscul              GVCAFVKVQLKDGSYHEDVGYGVSEGLRSKALSLEKARKEAVTDGLKRALRSFGNALGNC 166
                       *  *:*:* * .*:::**:***. *. : *. ::*:*:*.****.***:**.********

S. cerevi              LYDKDFLAKIDKVKFDPPDFDENNLFRPTDEISESSRTNTLHENQEQQQYPNKRRQLTKV 240
K. lactis              LYDKDFLAKIDKVKFDPPDFDEGNLFRPADELSEMSRSNMVGDAHTEGPSLKKRSLTNED 200
H. sapien              ILDKDYLRSLNKLPRQLPLEVDLTKAKRQDLEPSVEEARYNSCRPNMALGHPQLQQVTSP 225
M. muscul              ILDKDYLRSLNKLPRQLPLDVDLTKTKREDFEPSVEQARYNSCRQNEALGLPKPQEVTSP 226
                       : ***:* .::*:  : *   : .  :  *  .. ..:.              :    ..

S. cerevi              TNTNPDSTKNLVKIEN--TVSRGTPMMAAPAEANSKNSSNKDTDLKSLDASKQDQDDLLD 298
K. lactis              RNAVPSPPAQQTYRSNNHTTQKRAPIQAVTASASPNEETSN---------QQQDPDDLLD 251
H. sapien              SRPS-----HAVIPADQDCSSRSLSSSAVESEATHQRKLR-------------------- 260
M. muscul              CRSSPPHDSNIKLQGAKDISSSCSLAATLESDATHQRKLRK------------------- 267
                        ..      :      .   .      :  :.*. :..

S. cerevi              DSLMFSDDFQDDDLINMGNTNSNVLTTEKDPVVAKQSPTASSNPEAEQITFVTAKAATSV 358
K. lactis              DSFMFSDEIQDDDLLNMNTTTNNKNSTNSSTTTTTISDEATG--IISPVTFVTAKAATSL 309
H. sapien              -----QKQLQQQFRERME--KQQVRVSTPSAEKSEAAPPAPPVTHSTPVTVSEPLLEKDF 313
M. muscul              ---LRQKQLQQQFREQMETRRQSHAPAEEVAAKHAAVLPAPP-KHSTPVTAASELLQEKV 323
                            ..::*::   .*    ..   .       *.      :*       ..

S. cerevi              QNERYIGEESIFDPKYQAQSIRHTVDQTTSKHIPASVLKDKTMTTARDSVYEKFAPKGKQ 418
K. lactis              QHKDPIPSGSMFDPKFQAQSIRHTVDQSVSTPVRATILKEKGLDSDRSSIYSKFAPKGKE 369
H. sapien              LAGVTQELIKTLEDNSEKWAVTPDAGDGVVKPSSRADPAQTSDTLALNNQMVTQNRTPHS 373
M. muscul              VF------PDNLEENLEMWDLTPDLED-IIKPLCRGEPAQTSATRTFNN----QDSVPHI 372
                             . :: : :   :     :   .        :.      ..   .     :

S. cerevi              LSMKNNDKELGPHMLEGAGNQVPRETTPIKTNATAFPPAAAPRFAPPSKVVHPNGNGAVP 478
K. lactis              LSGTTTNSEPYVAAPQTSATESNRS-TPTRSNAQLAGPQPAPQLQGPQ------------ 416
H. sapien              VCHQKPQAKSGSWDLQTYSADQRTTGNWES--HRKSQDMKKRKYDPS------------- 418
M. muscul              HCHQKPQEKPGPGHLQTCNTNQHVLGSRDSEPHRKSQDLKKRKLDPS------------- 419
                        .  . : :      :    :     .               :   .

S. cerevi              AVPQQRSTRREVGRPKINPLHARKPT- 504
K. lactis              --------RTQLGRPRMLQQPNRRNVS 435
H. sapien              --------------------------- 
M. muscul              ---------------------------
```

Sequence alignment of Rad52 using ClustalW. From the top: *S. cerevisiae* Rad52, *K. lactis* Rad52, *H. sapiens* Rad52, and *M. musculus* Rad52. The colour code is based on the amino acids charges, hydrophobicity and polarity. In red is non-polar, hydrophobic residues, in green polar, hydrophilic residues with no net charge (except positively charged histidine). In blue negatively charged, polar and hydrophilic residues and in pink positively charged, polar and hydrophilic residues.

6.2 PSORT II – classification algorithm

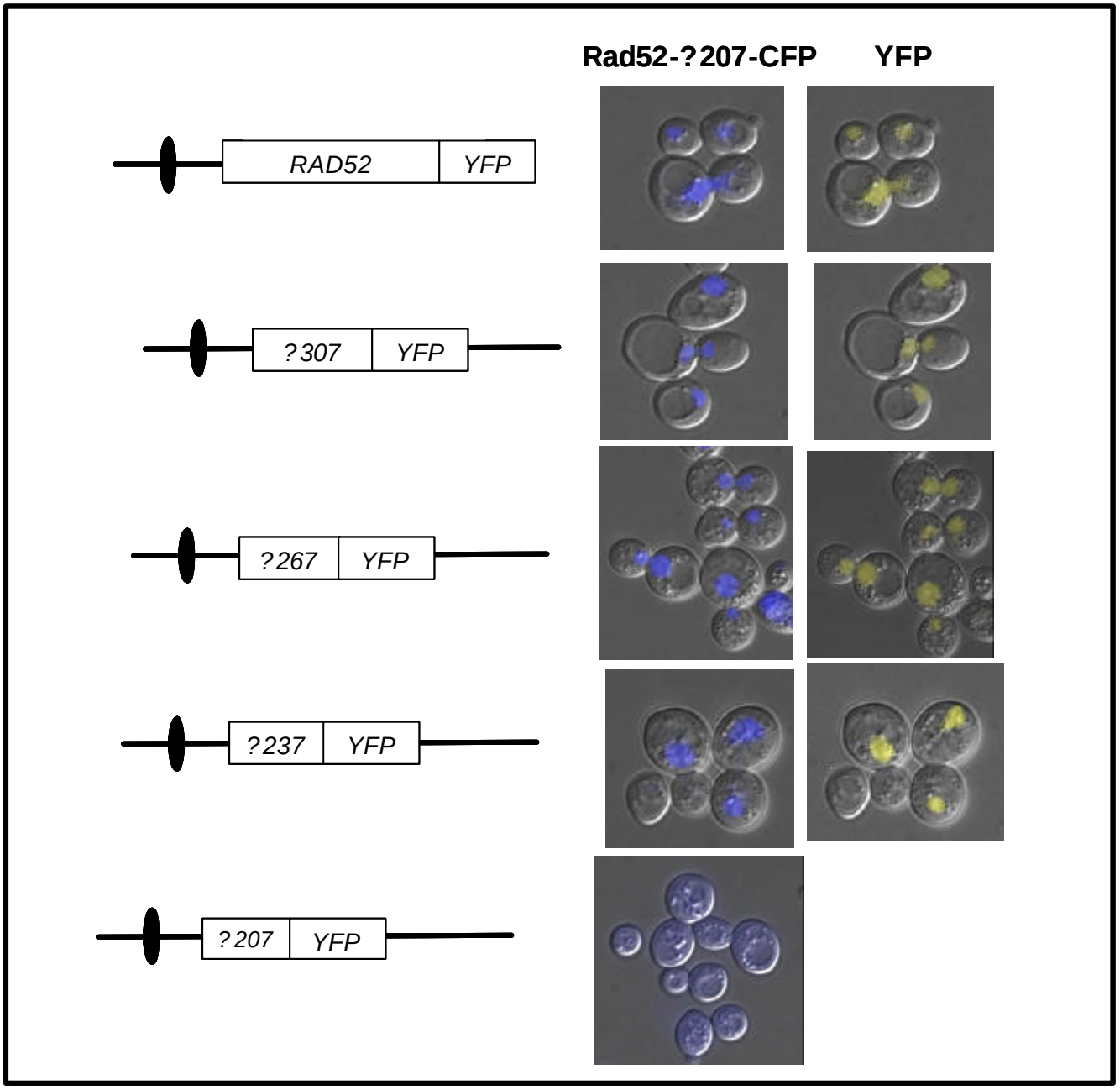
PSORT is a program for predicting subcellular localization by recognizing sorting signals in protein sequences (Nakai and Horton, 1999). The updated version of the PSORT program, PSORT II, is based on the classification algorithm *k*-nearest neighbours (KNN).
The KNN-classifier is a pattern recognition algorithm which assigns a label (or class) to a data instance based on that data instance's similarity to other instances which has already been 'labelled'. The algorithm chooses the k data instances ('neighbours') in the training set which has highest similarity to the new and unlabelled instance. Thereafter it is being assigned the class which dominates that k nearest neighbours. The similarity metric that instances are compared by can vary according to the classification task but euclidian distance (basic numerical difference) is the simplest and most widely used.

In this setting the new data instance is a query protein sequence and its label or class that we are trying to predict is its cellular localization. The algorithm works by first comparing the query sequence to a number of predefined sequence features that are known to determine cellular localization. A 'hit-score' for each of these features is calculated for the query and this set of feature scores is then compared to the corresponding scores of the training set using euclidian distance. The class that matches most of the k nearest neighbours to the query is assigned as the most likely localization. The number of neighbours has been set by the authors to 21 as this number gave the best prediction performance in tests.

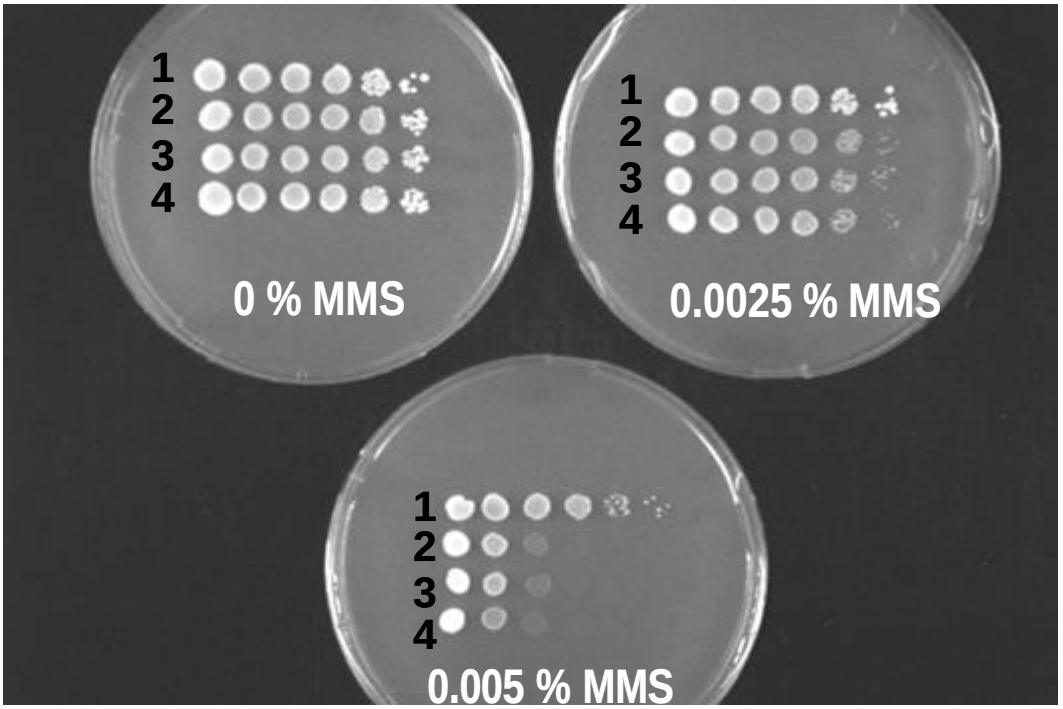
The query is for instance a protein X of unknown cellular location and the category is the cellular compartments. A dataset of 1484 yeast protein sequences with known classes (10 different localization sites) are used as training set.
The PSORT II classifier predicted ScRad52 to locate in the nucleus. The number of proteins that are incorrectly predicted are mostly due to confusing cytoplasmic proteins with nuclear proteins and *vice versa,* which is ascribed to the fundamental difficulty in identifying nuclear proteins (Nakai and Horton, 1999). The nuclear proteins are difficult to classify because the nuclear localization signals might be bipartite as mentioned above or the transport might be mediated by interaction with an NLS-containing protein partner.

6.3 Co-localization study

Rad52-Δ207-CFP is translocated to the nucleus when co-expressed with *RAD52-YFP*, *rad52-D307-YFP*, *rad52-D267-YFP* or *rad52-D237-YFP*.



6.4 Complementation assay using YFP-tagged *RAD52* mutant alleles



A *rad52-D327* strain is not complemented when transformed with YFP-tagged mutant alleles *rad52-MC-YFP-NLS* (2) or *rad52-C-YFP-NLS* (3). Wild type *RAD52-YFP* (1) show only mild MMS sensitivity even at 0.005 % MMS, whereas the strain transformed with an empty vector (4) are highly sensitive to the MMS induced breaks.

Reference List

Aboussekhra,A., Chanet,R., Adjiri,A., and Fabre,F. (1992). Semidominant suppressors of Srs2 helicase mutations of Saccharomyces cerevisiae map in the RAD51 gene, whose sequence predicts a protein with similarities to procaryotic RecA proteins. Mol. Cell Biol. *12*, 3224-3234.

Adzuma,K., Ogawa,T., and Ogawa,H. (1984). Primary structure of the RAD52 gene in Saccharomyces cerevisiae. Mol. Cell Biol. *4*, 2735-2744.

Alani,E., Subbiah,S., and Kleckner,N. (1989). The yeast RAD50 gene encodes a predicted 153-kD protein containing a purine nucleotide-binding domain and two large heptad-repeat regions. Genetics *122*, 47-57.

Alani,E., Thresher,R., Griffith,J.D., and Kolodner,R.D. (1992). Characterization of DNA-binding and strand-exchange stimulation properties of y-RPA, a yeast single-strand-DNA-binding protein. J. Mol. Biol. *227*, 54-71.

Anderson,D.E., Trujillo,K.M., Sung,P., and Erickson,H.P. (2001). Structure of the Rad50 x Mre11 DNA repair complex from Saccharomyces cerevisiae by electron microscopy. J. Biol. Chem. *276*, 37027-37033.

Arai,N., Ito,D., Inoue,T., Shibata,T., and Takahashi,H. (2005). Heteroduplex joint formation by a stoichiometric complex of Rad51 and Rad52 of Saccharomyces cerevisiae. J. Biol. Chem. *280*, 32218-32229.

Asleson,E.N., Okagaki,R.J., and Livingston,D.M. (1999). A core activity associated with the N terminus of the yeast RAD52 protein is revealed by RAD51 overexpression suppression of C-terminal rad52 truncation alleles. Genetics *153*, 681-692.

Aylon,Y. and Kupiec,M. (2004). DSB repair: the yeast paradigm. DNA Repair (Amst) *3*, 797-815.

Bai,Y., Davis,A.P., and Symington,L.S. (1999). A novel allele of RAD52 that causes severe DNA repair and recombination deficiencies only in the absence of RAD51 or RAD59. Genetics *153*, 1117-1130.

Bai,Y. and Symington,L.S. (1996). A Rad52 homolog is required for RAD51-independent mitotic recombination in Saccharomyces cerevisiae. Genes Dev. *10*, 2025-2037.

Basile,G., Aker,M., and Mortimer,R.K. (1992). Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene RAD51. Mol. Cell Biol. *12*, 3235-3246.

Baumann,P., Benson,F.E., and West,S.C. (1996). Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions in vitro. Cell *87*, 757-766.

Bendixen,C., Sunjevaric,I., Bauchwitz,R., and Rothstein,R. (1994). Identification of a mouse homologue of the Saccharomyces cerevisiae recombination and repair gene, RAD52. Genomics *23*, 300-303.

Bezzubova,O.Y., Schmidt,H., Ostermann,K., Heyer,W.D., and Buerstedde,J.M. (1993). Identification of a chicken RAD52 homologue suggests conservation of the RAD52 recombination pathway throughout the evolution of higher eukaryotes. Nucleic Acids Res. *21*, 5945-5949.

Bi,B., Rybalchenko,N., Golub,E.I., and Radding,C.M. (2004). Human and yeast Rad52 proteins promote DNA strand exchange. Proc. Natl. Acad. Sci. U. S. A.

Bochkareva,E. *et al.* (2005). Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. Proc. Natl. Acad. Sci. U. S. A *102*, 15412-15417.

Boulikas,T. (1997). Nuclear import of DNA repair proteins. Anticancer Res. *17*, 843-863.

Boulikas,T. (1994). Putative nuclear localization signals (NLS) in protein transcription factors. J. Cell Biochem. *55*, 32-58.

Boundy-Mills,K.L. and Livingston,D.M. (1993). A Saccharomyces cerevisiae RAD52 allele expressing a C-terminal truncation protein: activities and intragenic complementation of missense mutations. Genetics *133*, 39-49.

Brill,S.J. and Stillman,B. (1989). Yeast replication factor-A functions in the unwinding of the SV40 origin of DNA replication. Nature *342*, 92-95.

Cedano,J., Aloy,P., Perez-Pons,J.A., and Querol,E. (1997). Relation between amino acid composition and cellular location of proteins. J. Mol. Biol. *266*, 594-600.

Chen,L., Trujillo,K., Ramos,W., Sung,P., and Tomkinson,A.E. (2001). Promotion of Dnl4-catalyzed DNA end-joining by the Rad50/Mre11/Xrs2 and Hdf1/Hdf2 complexes. Mol. Cell *8*, 1105-1115.

Chou K-C and Elrod D.W (1999). Protein subcellular location prediction. Protein Engineering *12*, 107-118.

Clerici,M., Mantiero,D., Lucchini,G., and Longhese,M.P. (2005). The Saccharomyces cerevisiae Sae2 Protein Promotes Resection and Bridging of Double Strand Break Ends. J. Biol. Chem. *280*, 38631-38638.

Clever,B., Schmuckli-Maurer,J., Sigrist,M., Glassner,B.J., and Heyer,W.D. (1999). Specific negative effects resulting from elevated levels of the recombinational repair protein Rad54p in Saccharomyces cerevisiae. Yeast *15*, 721-740.

Cortes-Ledesma,F., Malagon,F., and Aguilera,A. (2004). A novel yeast mutation, rad52-L89F, causes a specific defect in Rad51-independent recombination that correlates with a reduced ability of Rad52-L89F to interact with Rad59. Genetics *168*, 553-557.

Davis,A.P. and Symington,L.S. (2001). The yeast recombinational repair protein Rad59 interacts with Rad52 and stimulates single-strand annealing. Genetics *159*, 515-525.

Davis,A.P. and Symington,L.S. (2003). The Rad52-Rad59 complex interacts with Rad51 and replication protein A. DNA Repair (Amst) *2*, 1127-1134.

de,J.M., van,N.J., van,G., Dekker,C., Kanaar,R., and Wyman,C. (2001). Human Rad50/Mre11 is a flexible complex that can tether DNA ends. Mol. Cell *8*, 1129-1135.

Dean,F.B., Borowiec,J.A., Eki,T., and Hurwitz,J. (1992). The simian virus 40 T antigen double hexamer assembles around the DNA at the replication origin. J. Biol. Chem. *267*, 14129-14137.

Dingwall,C. and Laskey,R.A. (1991). Nuclear targeting sequences--a consensus? Trends Biochem. Sci. *16*, 478-481.

Donovan,J.W., Milne,G.T., and Weaver,D.T. (1994). Homotypic and heterotypic protein associations control Rad51 function in double-strand break repair. Genes Dev. *8*, 2552-2562.

Dornfeld,K.J. and Livingston,D.M. (1991). Effects of controlled RAD52 expression on repair and recombination in Saccharomyces cerevisiae. Mol. Cell Biol. *11*, 2013-2017.

Erdeniz,N., Mortensen,U.H., and Rothstein,R. (1997). Cloning-free PCR-based allele replacement methods. Genome Res. *7*, 1174-1183.

Erdile,L.F., Heyer,W.D., Kolodner,R., and Kelly,T.J. (1991). Characterization of a cDNA encoding the 70-kDa single-stranded DNA-binding subunit of human replication protein A and the role of the protein in DNA replication. J. Biol. Chem. *266*, 12090-12098.

Essers,J., Houtsmuller,A.B., van Veelen,L., Paulusma,C., Nigg,A.L., Pastink,A., Vermeulen,W., Hoeijmakers,J.H., and Kanaar,R. (2002). Nuclear dynamics of RAD52 group homologous recombination proteins in response to DNA damage. EMBO J *21*, 2030-2037.

Fahrenkrog,B. and Aebi,U. (2003). The nuclear pore complex: nucleocytoplasmic transport and beyond. Nat. Rev. Mol. Cell Biol. *4*, 757-766.

Fairman,M.P. and Stillman,B. (1988). Cellular factors required for multiple stages of SV40 DNA replication in vitro. EMBO J. *7*, 1211-1218.

Fan,H.Y., Cheng,K.K., and Klein,H.L. (1996). Mutations in the RNA polymerase II transcription machinery suppress the hyperrecombination mutant hpr1 delta of Saccharomyces cerevisiae. Genetics *142*, 749-759.

Firmenich,A.A., Elias-Arnanz,M., and Berg,P. (1995). A novel allele of Saccharomyces cerevisiae RFA1 that is deficient in recombination and repair and suppressible by RAD52. Mol. Cell Biol. *15*, 1620-1631.

Fortin,G.S. and Symington,L.S. (2002). Mutations in yeast Rad51 that partially bypass the requirement for Rad55 and Rad57 in DNA repair by increasing the stability of Rad51-DNA complexes. EMBO J. *21*, 3160-3170.

Fujimori,A., Tachiiri,S., Sonoda,E., Thompson,L.H., Dhar,P.K., Hiraoka,M., Takeda,S., Zhang,Y., Reth,M., and Takata,M. (2001). Rad52 partially substitutes for the Rad51 paralog XRCC3 in maintaining chromosomal integrity in vertebrate cells. EMBO J *20*, 5513-5520.

Game,J.C. and Mortimer,R.K. (1974). A genetic study of x-ray sensitive mutants in yeast. Mutat. Res. *24*, 281-292.

Gasior,S.L., Wong,A.K., Kora,Y., Shinohara,A., and Bishop,D.K. (1998). Rad52 associates with RPA and functions with rad55 and rad57 to assemble meiotic recombination complexes. Genes Dev. *12*, 2208-2221.

Gorlich,D. and Mattaj,I.W. (1996). Nucleocytoplasmic transport. Science *271*, 1513-1518.

Haaf,T., Golub,E.I., Reddy,G., Radding,C.M., and Ward,D.C. (1995). Nuclear foci of mammalian Rad51 recombination protein in somatic cells after DNA damage and its localization in synaptonemal complexes. Proc. Natl. Acad. Sci. U. S. A *92*, 2298-2302.

Hays,S.L., Firmenich,A.A., and Berg,P. (1995). Complex formation in yeast double-strand break repair: participation of Rad51, Rad52, Rad55, and Rad57 proteins. Proc. Natl. Acad. Sci. U. S. A *92*, 6925-6929.

Hays,S.L., Firmenich,A.A., Massey,P., Banerjee,R., and Berg,P. (1998). Studies of the interaction between Rad52 protein and the yeast single- stranded DNA binding protein RPA. Mol. Cell Biol. *18*, 4400-4406.

Hefferin,M.L. and Tomkinson,A.E. (2005). Mechanism of DNA double-strand break repair by non-homologous end joining. DNA Repair (Amst) *4*, 639-648.

Heim,R. and Tsien,R.Y. (1996). Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. Curr. Biol. *6*, 178-182.

Heyer,W.D., Rao,M.R., Erdile,L.F., Kelly,T.J., and Kolodner,R.D. (1990). An essential Saccharomyces cerevisiae single-stranded DNA binding protein is homologous to the large subunit of human RP-A. EMBO J. *9*, 2321-2329.

Hicks,G.R. and Raikhel,N.V. (1995). Protein import into the nucleus: an integrated view. Annu. Rev. Cell Dev. Biol. *11*, 155-188.

Horton,P. and Nakai,K. (1996). A probabilistic classification system for predicting the cellular localization sites of proteins. Proc. Int. Conf. Intell. Syst. Mol. Biol. *4*, 109-115.

Ivanov,E.L., Sugawara,N., White,C.I., Fabre,F., and Haber,J.E. (1994). Mutations in XRS2 and RAD50 delay but do not prevent mating-type switching in Saccharomyces cerevisiae. Mol. Cell Biol. *14*, 3414-3425.

Jablonovich,Z., Liefshitz,B., Steinlauf,R., and Kupiec,M. (1999). Characterization of the role played by the RAD59 gene of Saccharomyces cerevisiae in ectopic recombination. Curr. Genet. *36*, 13-20.

Jackson,D., Dhar,K., Wahl,J.K., Wold,M.S., and Borgstahl,G.E. (2002). Analysis of the human replication protein A:Rad52 complex: evidence for crosstalk between RPA32, RPA70, Rad52 and DNA. J Mol. Biol. *321*, 133-148.

Jackson,S.P. (2002). Sensing and repairing DNA double-strand breaks. Carcinogenesis *23*, 687-696.

Jasin,M. (2002). Homologous repair of DNA damage and tumorigenesis: the BRCA connection. Oncogene *21*, 8981-8993.

Jiang,H., Xie,Y., Houston,P., Stemke-Hale,K., Mortensen,U.H., Rothstein,R., and Kodadek,T. (1996). Direct association between the yeast Rad51 and Rad54 recombination proteins. J. Biol. Chem. *271*, 33181-33186.

Johnson,R.D. and Symington,L.S. (1995). Functional differences and interactions among the putative RecA homologs Rad51, Rad55, and Rad57. Mol. Cell Biol. *15*, 4843-4850.

Johzuka,K. and Ogawa,H. (1995). Interaction of Mre11 and Rad50: two proteins required for DNA repair and meiosis-specific double-strand break formation in Saccharomyces cerevisiae. Genetics *139*, 1521-1532.

Kagawa,W., Kurumizaka,H., Ikawa,S., Yokoyama,S., and Shibata,T. (2001). Homologous pairing promoted by the human Rad52 protein. J Biol. Chem *276*, 35201-35208.

Kagawa,W., Kurumizaka,H., Ishitani,R., Fukai,S., Nureki,O., Shibata,T., and Yokoyama,S. (2002). Crystal structure of the homologous-pairing domain from the human Rad52 recombinase in the undecameric form. Mol. Cell *10*, 359-371.

Kalderon,D., Richardson,W.D., Markham,A.F., and Smith,A.E. (1984a). Sequence requirements for nuclear location of simian virus 40 large-T antigen. Nature *311*, 33-38.

Kalderon,D., Roberts,B.L., Richardson,W.D., and Smith,A.E. (1984b). A short amino acid sequence able to specify nuclear location. Cell *39*, 499-509.

Kalderon,D. and Smith,A.E. (1984). In vitro mutagenesis of a putative DNA binding domain of SV40 large-T. Virology *139*, 109-137.

Kantake,N., Sugiyama,T., Kolodner,R.D., and Kowalczykowski,S.C. (2003). The recombination-deficient mutant RPA (rfa1-t11) is displaced slowly from single-stranded DNA by Rad51 protein. J. Biol. Chem. *278*, 23410-23417.

Kiers,J., Zeeman,A.M., Luttik,M., Thiele,C., Castrillo,J.I., Steensma,H.Y., van Dijken,J.P., and Pronk,J.T. (1998). Regulation of alcoholic fermentation in batch and chemostat cultures of Kluyveromyces lactis CBS 2359. Yeast *14*, 459-469.

Krejci,L., Song,B., Bussen,W., Rothstein,R., Mortensen,U.H., and Sung,P. (2002). Interaction with Rad51 is indispensable for recombination mediator function of Rad52. J Biol. Chem *277*, 40132-40141.

Krejci,L., Van,K.S., Li,Y., Villemain,J., Reddy,M.S., Klein,H., Ellenberger,T., and Sung,P. (2003). DNA helicase Srs2 disrupts the Rad51 presynaptic filament. Nature *423*, 305-309.

Krogh,B.O., Llorente,B., Lam,A., and Symington,L.S. (2005). Mutations in Mre11 phosphoesterase motif I that impair Saccharomyces cerevisiae Mre11-Rad50-Xrs2 complex stability in addition to nuclease activity. Genetics.

Kumar,J.K. and Gupta,R.C. (2004). Strand exchange activity of human recombination protein Rad52. Proc. Natl. Acad. Sci. U. S. A.

Lewis,L.K., Storici,F., Van Komen,S., Calero,S., Sung,P., and Resnick,M.A. (2004). Role of the nuclease activity of Saccharomyces cerevisiae Mre11 in repair of DNA double-strand breaks in mitotic cells. Genetics *166*, 1701-1713.

Lichten,M. (2005). Rad50 connects by hook or by crook. Nat. Struct. Mol. Biol. *12*, 392-393.

Lim,D.S. and Hasty,P. (1996). A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. Mol. Cell Biol. *16*, 7133-7143.

Lisby,M., Barlow,J.H., Burgess,R.C., and Rothstein,R. (2004). Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. Cell *118*, 699-713.

Lisby,M., Mortensen,U.H., and Rothstein,R. (2003a). Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. Nat. Cell Biol. *5*, 572-577.

Lisby,M., ntunez de,M.A., Mortensen,U.H., and Rothstein,R. (2003b). Cell cycle-regulated centers of DNA double-strand break repair. Cell Cycle *2*, 479-483.

Lisby,M., Rothstein,R., and Mortensen,U.H. (2001). Rad52 forms DNA repair and recombination centers during S phase. Proc. Natl. Acad. Sci. U. S. A *98*, 8276-8282.

Liu,Y., Li,M., Lee,E.Y., and Maizels,N. (1999). Localization and dynamic relocalization of mammalian Rad52 during the cell cycle and in response to DNA damage. Curr. Biol. *9*, 975-978.

Macara,I.G. (2001). Transport into and out of the nucleus. Microbiol. Mol. Biol. Rev. *65*, 570-94, table.

Mazin,A.V., Alexeev,A.A., and Kowalczykowski,S.C. (2003). A novel function of Rad54 protein. Stabilization of the Rad51 nucleoprotein filament. J. Biol. Chem. *278*, 14029-14036.

McKee,A.H. and Kleckner,N. (1997). A general method for identifying recessive diploid-specific mutations in Saccharomyces cerevisiae, its application to the isolation of mutants

blocked at intermediate stages of meiotic prophase and characterization of a new gene SAE2. Genetics *146*, 797-816.

Milne,G.T. and Weaver,D.T. (1993). Dominant negative alleles of RAD52 reveal a DNA repair/recombination complex including Rad51 and Rad52. Genes Dev. *7*, 1755-1765.

Miyazaki,T., Bressan,D.A., Shinohara,M., Haber,J.E., and Shinohara,A. (2004). In vivo assembly and disassembly of Rad51 and Rad52 complexes during double-strand break repair. EMBO J.

Moreno-Herrero,F., de Jager,M., Dekker,N.H., Kanaar,R., Wyman,C., and Dekker,C. (2005). Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. Nature *437*, 440-443.

Morita,T., Yoshimura,Y., Yamamoto,A., Murata,K., Mori,M., Yamamoto,H., and Matsushiro,A. (1993). A mouse homolog of the Escherichia coli recA and Saccharomyces cerevisiae RAD51 genes. Proc. Natl. Acad. Sci. U. S. A *90*, 6577-6580.

Mortensen,U.H., Bendixen,C., Sunjevaric,I., and Rothstein,R. (1996). DNA strand annealing is promoted by the yeast Rad52 protein. Proc. Natl. Acad. Sci. U. S. A *93*, 10729-10734.

Mortensen,U.H., Erdeniz,N., Feng,Q., and Rothstein,R. (2002). A molecular genetic dissection of the evolutionarily conserved N terminus of yeast Rad52. Genetics *161*, 549-562.

Moynahan,M.E. (2002). The cancer connection: BRCA1 and BRCA2 tumor suppression in mice and humans. Oncogene *21*, 8994-9007.

Muris,D.F. *et al.* (1994). Cloning of human and mouse genes homologous to RAD52, a yeast gene involved in DNA repair and recombination. Mutat. Res. *315*, 295-305.

Nakai,K. and Horton,P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem. Sci. *24*, 34-36.

Nakai,K. and Kanehisa,M. (1992). A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics *14*, 897-911.

Nakielny,S. and Dreyfuss,G. (1999). Transport of proteins and RNAs in and out of the nucleus. Cell *99*, 677-690.

Navadgi,V.M., Shukla,A., and Rao,B.J. (2005). Effect of DNA sequence and nucleotide cofactors on hRad51 binding to ssDNA: role of hRad52 in recruitment. Biochem. Biophys. Res. Commun. *334*, 696-701.

New,J.H., Sugiyama,T., Zaitseva,E., and Kowalczykowski,S.C. (1998). Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. Nature *391*, 407-410.

Ogawa,T., Yu,X., Shinohara,A., and Egelman,E.H. (1993). Similarity of the yeast RAD51 filament to the bacterial RecA filament. Science *259*, 1896-1899.

Ormo,M., Cubitt,A.B., Kallio,K., Gross,L.A., Tsien,R.Y., and Remington,S.J. (1996). Crystal structure of the Aequorea victoria green fluorescent protein. Science *273*, 1392-1395.

Paques,F. and Haber,J.E. (1999). Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. Microbiol. Mol. Biol. Rev. *63*, 349-404.

Park,M.S. (1995). Expression of human RAD52 confers resistance to ionizing radiation in mammalian cells. J. Biol. Chem. *270*, 15467-15470.

Park,M.S., Ludwig,D.L., Stigger,E., and Lee,S.H. (1996). Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. J Biol. Chem *271*, 18996-19000.

Passy,S.I., Yu,X., Li,Z., Radding,C.M., Masson,J.Y., West,S.C., and Egelman,E.H. (1999). Human Dmc1 protein binds DNA as an octameric ring. Proc. Natl. Acad. Sci. U. S. A *96*, 10684-10688.

Paull,T.T. and Gellert,M. (1999). Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex. Genes Dev. *13*, 1276-1288.

Petukhova,G., Stratton,S.A., and Sung,P. (1999). Single strand DNA binding and annealing activities in the yeast recombination factor Rad59. J. Biol. Chem. *274*, 33839-33842.

Petukhova,G., Sung,P., and Klein,H. (2000). Promotion of Rad51-dependent D-loop formation by yeast recombination factor Rdh54/Tid1. Genes Dev. *14*, 2206-2215.

Pfuetzner,R.A., Bochkarev,A., Frappier,L., and Edwards,A.M. (1997). Replication protein A. Characterization and crystallization of the DNA binding domain. J. Biol. Chem. *272*, 430-434.

Prakash,L. and Prakash,S. (1977a). Isolation and characterization of MMS-sensitive mutants of Saccharomyces cerevisiae. Genetics *86*, 33-55.

Prakash,S. and Prakash,L. (1977b). Increased spontaneous mitotic segregation in MMS-sensitive mutants of Saccharomyces cerevisiae. Genetics *87*, 229-236.

Ranatunga,W., Jackson,D., Lloyd,J.A., Forget,A.L., Knight,K.L., and Borgstahl,G.E. (2001). Human RAD52 exhibits two modes of self-association. J Biol. Chem *276*, 15876-15880.

Resnick,M.A. (1969). Genetic control of radiation sensitivity in Saccharomyces cerevisiae. Genetics *62*, 519-531.

Ristic,D., Modesti,M., Kanaar,R., and Wyman,C. (2003). Rad52 and Ku bind to different DNA structures produced early in double-strand break repair. Nucleic Acids Res. *31*, 5229-5237.

Sakuraba,Y., Schroeder,A.L., Ishii,C., and Inoue,H. (2000). A Neurospora double-strand-break repair gene, mus-11, encodes a RAD52 homologue and is inducible by mutagens. Mol. Gen. Genet. *264*, 392-401.

Shalguev,V.I., Kaboev,O.K., Sizova,I.A., Hegemann,P., and Lantsov,V.A. (2005). [Identification of Rad51 protein from Chlamydomonas reinhardtii: recombinational characteristics]. Mol. Biol. (Mosk) *39*, 112-119.

Shen,Z., Cloud,K.G., Chen,D.J., and Park,M.S. (1996a). Specific interactions between the human RAD51 and RAD52 proteins. J Biol. Chem *271*, 148-152.

Shen,Z., Denison,K., Lobb,R., Gatewood,J.M., and Chen,D.J. (1995). The human and mouse homologs of the yeast RAD52 gene: cDNA cloning, sequence analysis, assignment to human chromosome 12p12.2-p13, and mRNA expression in mouse tissues. Genomics *25*, 199-206.

Shen,Z., Peterson,S.R., Comeaux,J.C., Zastrow,D., Moyzis,R.K., Bradbury,E.M., and Chen,D.J. (1996b). Self-association of human RAD52 protein. Mutat. Res. *364*, 81-89.

Sherman F,F.G.H.J. (1986). Methods in Yeast Genetics. Cold Spring Harbor Lab. Press.

Shinohara,A., Ogawa,H., and Ogawa,T. (1992). Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. Cell *69*, 457-470.

Shinohara,A. and Ogawa,T. (1998). Stimulation by Rad52 of yeast Rad51-mediated recombination. Nature *391*, 404-407.

Shinohara,A., Shinohara,M., Ohta,T., Matsuda,S., and Ogawa,T. (1998). Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing. Genes Cells *3*, 145-156.

Singleton,M.R., Wentzell,L.M., Liu,Y., West,S.C., and Wigley,D.B. (2002). Structure of the single-strand annealing domain of human RAD52 protein. Proc. Natl. Acad. Sci. U. S. A *99*, 13492-13497.

Snow,R. (1967). Mutants of yeast sensitive to ultraviolet light. J. Bacteriol. *94*, 571-575.

Solinger,J.A. and Heyer,W.D. (2001). Rad54 protein stimulates the postsynaptic phase of Rad51 protein-mediated DNA strand exchange. Proc. Natl. Acad. Sci. U. S. A *98*, 8447-8453.

Solinger,J.A., Lutz,G., Sugiyama,T., Kowalczykowski,S.C., and Heyer,W.D. (2001). Rad54 protein stimulates heteroduplex DNA formation in the synaptic phase of DNA strand exchange via specific interactions with the presynaptic Rad51 nucleoprotein filament. J. Mol. Biol. *307*, 1207-1221.

Stasiak,A.Z., Larquet,E., Stasiak,A., Muller,S., Engel,A., Van Dyck,E., West,S.C., and Egelman,E.H. (2000). The human Rad52 protein exists as a heptameric ring. Curr. Biol. *10*, 337-340.

Sugawara,N., Ira,G., and Haber,J.E. (2000). DNA length dependence of the single-strand annealing pathway and the role of Saccharomyces cerevisiae RAD59 in double-strand break repair. Mol. Cell Biol. *20*, 5300-5309.

Sugawara,N., Wang,X., and Haber,J.E. (2003). In vivo roles of Rad52, Rad54, and Rad55 proteins in Rad51-mediated recombination. Mol. Cell *12*, 209-219.

Sugiyama,T. and Kowalczykowski,S.C. (2002). Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation. J. Biol. Chem. *277*, 31663-31672.

Sugiyama,T., New,J.H., and Kowalczykowski,S.C. (1998). DNA annealing by RAD52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. Proc. Natl. Acad. Sci. U. S. A *95*, 6049-6054.

Sugiyama,T., Zaitseva,E.M., and Kowalczykowski,S.C. (1997). A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the Saccharomyces cerevisiae Rad51 protein. J. Biol. Chem. *272*, 7940-7945.

Sung,P. (1994). Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. Science *265*, 1241-1243.

Sung,P. (1997b). Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein A to promote DNA strand exchange by Rad51 recombinase. Genes Dev. *11*, 1111-1121.

Sung,P. (1997a). Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. J Biol. Chem *272*, 28194-28197.

Sung,P., Krejci,L., Van,K.S., and Sehorn,M.G. (2003). Rad51 recombinase and recombination mediators. J. Biol. Chem. *278*, 42729-42732.

Sung,P. and Robberson,D.L. (1995). DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. Cell *82*, 453-461.

Sung,P., Trujillo,K.M., and Van Komen,S. (2000). Recombination factors of Saccharomyces cerevisiae. Mutat. Res. *451*, 257-275.

Symington,L.S. (2002). Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. Microbiol. Mol. Biol. Rev. *66*, 630-70, table.

Takahashi,N. and Dawid,I.B. (2005). Characterization of zebrafish Rad52 and replication protein A for oligonucleotide-mediated mutagenesis. Nucleic Acids Res. *33*, e120.

Tan,T.L., Essers,J., Citterio,E., Swagemakers,S.M., de,W.J., Benson,F.E., Hoeijmakers,J.H., and Kanaar,R. (1999). Mouse Rad54 affects DNA conformation and DNA-damage-induced Rad51 foci formation. Curr. Biol. *9*, 325-328.

Thomas,B.J. and Rothstein,R. (1989). Elevated recombination rates in transcriptionally active DNA. Cell *56*, 619-630.

Trujillo,K.M., Roh,D.H., Chen,L., Van,K.S., Tomkinson,A., and Sung,P. (2003). Yeast xrs2 binds DNA and helps target rad50 and mre11 to DNA ends. J. Biol. Chem. *278*, 48957-48964.

Trujillo,K.M. and Sung,P. (2001). DNA structure-specific nuclease activities in the Saccharomyces cerevisiae Rad50*Mre11 complex. J. Biol. Chem. *276*, 35458-35464.

Tsubouchi,H. and Ogawa,H. (1998). A novel mre11 mutation impairs processing of double-strand breaks of DNA during both mitosis and meiosis. Mol. Cell Biol. *18*, 260-268.

Usui,T., Ohta,T., Oshiumi,H., Tomizawa,J., Ogawa,H., and Ogawa,T. (1998). Complex formation and functional versatility of Mre11 of budding yeast in recombination. Cell *95*, 705-716.

van den,B.M., Vreeken,K., Zonneveld,J.B., Brandsma,J.A., Lombaerts,M., Murray,J.M., Lohman,P.H., and Pastink,A. (2001). Characterization of RAD52 homologs in the fission yeast Schizosaccharomyces pombe. Mutat. Res. *461*, 311-323.

Van Dyck,E., Stasiak,A.Z., Stasiak,A., and West,S.C. (2001). Visualization of recombination intermediates produced by RAD52-mediated single-strand annealing. EMBO Rep. *2*, 905-909.

van Veelen,L.R., Essers,J., van de Rakt,M.W., Odijk,H., Pastink,A., Zdzienicka,M.Z., Paulusma,C.C., and Kanaar,R. (2005). Ionizing radiation-induced foci formation of mammalian Rad51 and Rad54 depends on the Rad51 paralogs, but not on Rad52. Mutat. Res. *574*, 34-49.

Van,D.E., Hajibagheri,N.M., Stasiak,A., and West,S.C. (1998). Visualisation of human rad52 protein and its complexes with hRad51 and DNA. J. Mol. Biol. *284*, 1027-1038.

Van,D.E., Stasiak,A.Z., Stasiak,A., and West,S.C. (1999). Binding of double-strand breaks in DNA by human Rad52 protein. Nature *398*, 728-731.

Wold,M.S. (1997). Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu. Rev. Biochem. *66*, 61-92.

Wolner,B., Van Komen,S., Sung,P., and Peterson,C.L. (2003). Recruitment of the recombinational repair machinery to a DNA double-strand break in yeast. Mol. Cell *12*, 221-232.

Zou,H. and Rothstein,R. (1997). Holliday junctions accumulate in replication mutants via a RecA homolog-independent mechanism. Cell *90*, 87-96.